2017

G-Quadruplexes As Molecular Targets For Transcriptional Control

Rhianna Kyelle Morgan
University of Mississippi

G-QUADRUPLEXES AS MOLECULAR TARGETS FOR TRANSCRIPTIONAL CONTROL

A Dissertation
presented in fulfillment of requirements
for the degree of Doctorate of Philosophy
in the Department of BioMolecular Sciences
Division of Pharmacology
The University of Mississippi

by

RHIANNA K. MORGAN

March 2017

ABSTRACT

DNA G-quadruplexes (G4s) have gained interest in modern drug discovery efforts. The majority of G4 research, including the mass of this research, heavily focuses on cancer therapeutic development, as G4s are more pronounced in oncogenic promoters compared to the rest of the genome, our efforts extend towards neuroprotection as well. In particular, this work covers G4s in the promoters of MYC, VEGF, kRAS, and Bcl-2. MYC and VEGF harbor the two most well described promoter G4s to date, and their stabilization cultivates transcriptional silencing. Downregulation of MYC reduces cell growth and alters cell energetics; it is promising for a variety of cancer types, including lymphomas and leukemias. VEGF downregulation modulates angiogenesis, therapeutically benefiting most solid tumors. Lesser described G4s, kRAS and Bcl-2, act as transcriptional silencers and activators, respectively. kRAS is highly deregulated in 30% of cancers, and its downregulation can decrease tumor cell proliferation and survival. Decreases in Bcl-2 expression result in pro-apoptotic signaling. However, Bcl-2 $G4_1$ stabilization can cause transcriptional activation, and its upregulation reduces apoptosis, which is of therapeutic benefit for neurodegenerative diseases. The current study examines (a) the effect of epigenetic modification on VEGF G4 stability, (b) a variety of osmolytes and crowding co-solvents for their ability to recapitulate the physiological structure in single-stranded ex vivo conditions, (c) characterizing and targeting the biologically relevant G4 in the kRAS core promoter for a novel approach to targeting kRAS driven cancers, specifically pancreatic cancer, and (d) targeting one of four G4s ($G4_1$)

within the Bcl-2 promoter for neuroprotection. For all of the studies, significant changes in G4 structures, including loop directionality and number of competing isoforms, were examined by electronic circular dichroism. Electromobility shift assays differentiated inter- and intra-molecular structures, further distinguishing the distribution of isoforms. Other studies included chemical footprinting, FRET compound screening, luciferase, and in vitro cytotoxicity. This comprehensive understanding of the physiological conditions regulating G4 stability and function using co-solvents to examine the effects of small molecules ex vivo will best inform future drug discovery efforts on how to better predict active hits for in vitro evaluation for these and other promoters.

DEDICATION

This dissertation is dedicated to my beloved family, friends, and the BioMolecular Sciences Graduate Research Program.

LIST OF ABBREVIATIONS AND SYMBOLS

5’UTR	5’ Untranslated region
5-FAM	5-carboxyfluorescein
5-hmC	5-hydroxymethylcytosine
5-mC	5-methylcytosine
6-OHDA	6-hydroxydopamine
ADAM-15	A Disintegrin and metalloproteinase domain-containing protein 15
ALS	Amyotrophic lateral sclerosis (Lou Gehrig’s disease)
APS	Ammonium persulfate
BBB	Blood brain barrier
Bcl-2	B-cell lymphoma 2
CAT	Chloramphenicol acetyl transferase
ChIP	Chromatin immunoprecipitation
CNBP	Cellular nucleic acid-binding protein
COMT	Catechol-O-methyl transferase
DMEM	Dulbecco’s Modified Eagle Medium
DMS	Dimethyl sulfate
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
dsDNA	Double-stranded DNA
DTT	Dithiothreitol
ECD	Electronic circular dichroism
EDTA	Ethylenediaminetetraacetic acid
EMSA	Electromobility shift assay
EV	Empty vector plasmid

FBS	Fetal Bovine Serum
FRET	Fluorescence resonance energy transfer
G4	G-quadruplex
GAP	GTPase-activating protein
GAPDH	Glyceraldehyde 3-phosphate dehydrogenase
GDP	Guanosine diphosphate
GEF	Guanine nucleotide exchange factor
GTP	Guanosine-5'-triphosphate
H_2O_2	Hydrogen peroxide
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HIF-1	Hypoxia-inducible factor 1
hnRNPk	Heterogeneous nuclear ribonucleoprotein K
HRAS	Harvey rat sarcoma viral oncogene homolog
hTERT	Telomerase reverse transcriptase
IC_{50}	Half maximal inhibitory concentration
KO	Knockout
kRAS	Kirsten rat sarcoma viral oncogene
MAO-B	Monoamine oxidase B
MAPK	Mitogen-activated protein kinase
MAZ	MYC associated zinc finger protein
MeCN	Acetonitrile
MPP+	1-methyl-4-phenylpyridinium
MPPP	1-methyl-4-phenyl-4-propionoxypiperidine
MPTP	1-methyl-4-phenyl-1,2,3,6-tetrahydropyridine
MT	Mutant
MTS	3-(4,5-dimethylthiazol-2-yl)-5-(3-carboxymethoxyphenyl)-2-(4-sulfophenyl)-2H-tetrazolium
NCI	National Cancer Institute
NHE	Nuclease hypersensitivity element

NMR	Nuclear magnetic resonance spectroscopy
NRAS	Neuroblastoma rat sarcoma viral oncogene homolog
NSC	New synthetic compound
p53	Tumor protein 53
PAGE	Polyacrylamide gel electrophoresis
PCR	Polymerase chain reaction
pcv	Packed cell volume
PD	Parkinson's disease
PDGF-A	Platelet derived growth factor α
PEG	Polyethylene glycol
PMS	Phenazine methosulfate
PMSF	Phenylmethylsulfonyl fluoride
PNA	Peptide nucleic acid
pnv	Packed nuclear volume
pRB	Retinoblastoma protein
qRT-PCR	Quantitative real-time polymerase chain reaction
RNA	Ribonucleic acid
RNA Pol II	RNA polymerase II
RPMI	Roswell Park Memorial Institute
siRNA	Small interfering RNA
Sp1	Specificity protein 1
ssDNA	Single-stranded DNA
SVD	Singular value decomposition
TAMRA	Tetramethylrhodamine
TBE	Tris/Borate/EDTA buffer
TEMED	Tetramethylethylenediamine
TERRA	Telomeric repeat-containing RNA
T_M	Melting temperature
TSS	Transcriptional start site
VEGF	Vascular endothelial growth factor

WT1	Wilm’s tumor 1

ACKNOWLEDGEMENTS

I express my deepest appreciation to my mentor, Dr. Tracy Brooks, for her constant support and encouragement over the last few years. Her guidance and enthusiasm was always an inspiration to overcome research and personal troubles alike. Both her wisdom and trainings are invaluable for my future career. She will be a life-long friend.

Next, I would like to thank my committee members: Dr. Kristie Willett, Dr. John Rimoldi, and Dr. Christopher McCurdy, for their advice, support, and sometimes just a lending ear throughout my graduate studies.

I greatly appreciate the help and friendship from my colleagues: Harshul Batra, Taisen Hao, and Disha Prabhu, and all the undergraduates and pharmacy students who have helped with the projects presented herein.

I am also grateful to all of the faculty, staff, and graduate students in the Department of BioMolecular Sciences. I am fortunate to have been accepted into this program to initiate my journey in the world of research. The memories and experiences gained while at OleMiss are ones I will cherish my entire life.

Finally, I would like to thank my family for their endless love, tireless support, and push to pursue my dreams. Thank you for always rooting for me during the good times and the bad.

This work was supported by Dr. Tracy Brooks' startup funds from the University of Mississippi, National Institutes of Health (NIH) grant R15-CA173667-01A1, Department of Defense Peer Reviewed Cancer Research Program (DoD-PRCRP) grant CA130229, Rhianna Morgan's University of Mississippi Graduate Student Council (GSC) Research Grant in 2014, and the Institutional Development Award (IDeA) grant P20GM104932 from the National Institute of General Medical Sciences (NIGMS).

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1. INTRODUCTION

Non-canonical DNA structures, such as G-quadruplexes (G4s), show promise as molecular targets for therapeutics, specifically in oncology. Unlike typical B-DNA, G4s form within a single strand leaving the complementary strand to remain continuous. G4-forming regions are contained within very guanine dense sequences. This structure requires four guanines to hydrogen bond together, each acting as both proton donor and acceptor, to form a tetrad with the interconnecting nucleotides forming loops to connect guanines in the planes (Figure 1-1). This interaction between guanines is termed Hoogsteen base-pairing.

With an electronegative pocket at its center due to carbonyl oxygens, tetrads have increased stability when a monovalent cation is present (Williamson 1994; Guschlbauer et al. 1990). Many monovalent cations have been reported to stabilize tetrads, but the most commonly used are sodium and potassium. Early studies compared a manifold of cations ranging in ionic radii and found an inverse relationship between hydration energy and radii. Cations like K^+, Rb^+, Na^+, Cs^+, and Li^+ were linked to complex G4 stability, but out of these alkalines, K+ produces the most stable intramolecular G4 (Bryan & Baumann 2011).

G4s are polymorphic and can be categorized into four classes (Brooks & Hurley 2010). The simplest forms are contained within Class I. This is when one form predominates in a single G-rich region with differences only being observed within loop

arrangements. A prime example of a Class I G4 is MYC. In G4-forming regions that allow more than one structure to form and are separated by an extended region of approximately three helical turns are contained within Class II. These G4s are typically dissimilar from each other, as seen with c-KIT and kRAS. Class III is similar to the second class as two G4s form, but they occur in closer proximity. The best example of this type is hTERT where two G4s form with only 4-19 intervening bases, and have interstructural interactions. The final class resembles the first in that structure formation occurs within a single G-rich region of DNA, but this class differs as more than one isoform exists within the same G4-forming region. For instance, the Bcl-2 promoter contains of a 39 bp G4-forming region ($G4_1$) harboring six runs of guanines which allows at least three different G4s to form.

The first record of G4 formation was in 1910 when a Scandinavian chemist, Ivar Bang, discovered gel formation in aqueous solutions of high concentrations of guanosine derivatives (Bang 1910; Lagnado 2013). Interestingly, this phenomenon was described within fifty years of the identification of nuclein, or the nucleic acids (Miescher 1871; Wolf 2003; Pray 2008), however, it was not until 1962 that this mystery was solved following a further understanding of nucleic acids (Chargaff 1950) and after the DNA double helix model was proposed (Watson & Crick 1953). Fiber x-ray diffraction data portrayed “G-quartet” (tetrad) formation in telomeric DNA (Gellert et al. 1962). This sparked the G4 field of research, and the proposal that telomeric G4s could form and be therapeutic targets.

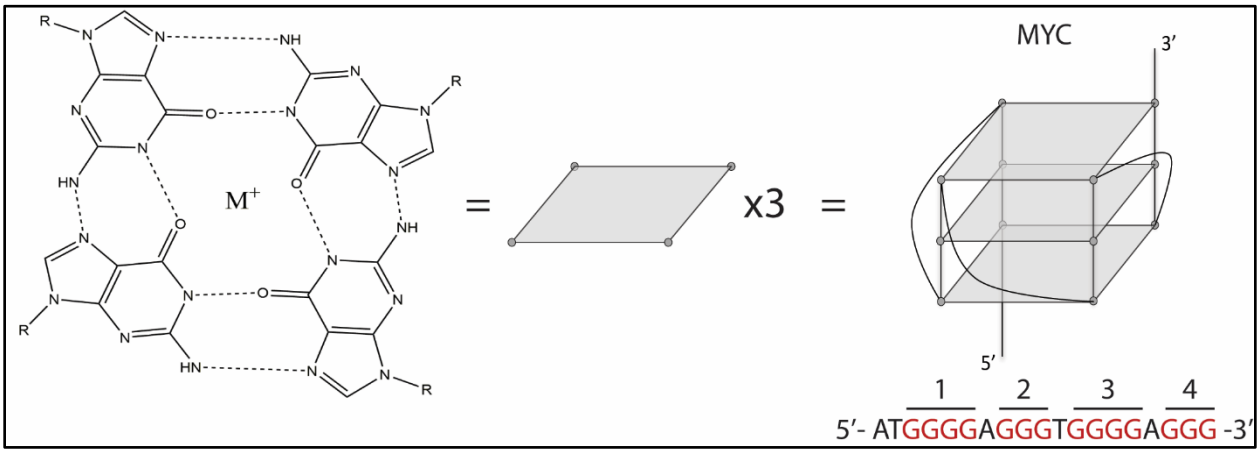


Figure 1-1. Guanine arrangement for G4 formation. In guanine rich sequences, four guanines Hoogsteen base pair to form a tetrad. Three or more tetrads stack on top of each other through π–π interactions, and the interconnecting nucleotides form the loops configured around the structure. The electronegative carbonyl oxygen pocket at the center of each tetrad allow for increased structural stability with a monovalent cation present. The structure described here is the initial promoter G4 elucidated within the MYC promoter (Adapted from Brooks & Hurley 2010).

Telomeres exist on the ends of eukaryotic chromosomes and are very dynamic, as they contain a single-stranded 3' overhang of approximately 200 nucleotides (Buscaglia et al. 2013). These DNA-protein structures can protect chromosome ends through a capping function in order to prevent DNA double-stranded breaks and from being processed by many other DNA damage pathways (Bidzinska et al. 2013). Telomeres shorten during cell division, and this shortening process is accelerated when telomere uncapping occurs. If telomeres become significantly short then the cell will no longer continue to proliferate and will instead enter senescence. However, there exists an enzyme, telomerase, that is expressed in fetal tissue, adult germ cell, and tumor cells and are responsible for chromosome lengthening and stabilization (Hanahan & Weinberg 2000; Hanahan & Weinberg 2011). Its upregulation results in the replicative immortality phenotype seen in many cancers. Telomeres harbor tandem repeats of guanine rich sequences (Henderson et al. 1990). Thus, the discovery of G4s forming within these regions came to no surprise.

Telomeric G4s gained interest as therapeutic targets for cancer by acting on or controlling telomerase which is upregulated in approximately 85% of cancers (Kim et al. 1994; Neidle & Parkinson 2003). Telomeric G4 caps block telomerase recruitment, resulting in cell senescence and apoptosis and acting to protect against extensive cellular division and mutation accumulation (Paeschke et al. 2008; Lipps & Rhodes 2009; Neidle 2009; Neidle 2010). Collectively, the characterization of telomeric structures gave further insight into overall G4 formation, more specifically the differences in loop configurations. Based on the initial telomeric G4 findings using denaturing polyacrylamide gel electrophoresis and NMR studies, there were several isoforms existing with differences

in the loops and in structure, including both intra- and intermolecular types (Huppert 2010). Briefly, intramolecular structures form within a single strand of DNA or RNA, and intermolecular G4s occur through the organization of two or more strands (Bryan & Baumann 2011). Intramolecular G4s are the structures of most interest as they are considered to be physiologically relevant with higher potential for small molecule targeting (Chen & Yang 2013). Also, they are polymorphic due to loop configurations. Loops can vary between G4s depending on the sequence which controls the final topology of the G4. Loops can link continuous guanines diagonally or laterally from the top or bottom of tetrads. A parallel conformation occurs when loops connect guanines located at the top to the bottom diagonally, using a double-chain reversal. Antiparallel structures contain loops connecting guanines in a diagonal or lateral fashion. There are G4s that exist in hybrid arrangements harboring both parallel and antiparallel loop directionalities known as mixed structures. The most stable G4s characterized have short loops in a parallel conformation as shown in Figure 1-1.

Outside DNA telomeric regions, G4s have been characterized in several gene promoters such as MYC, VEGF, c-KIT, hTERT, kRAS, and Bcl-2. These structures have the ability to activate or suppress transcription initiation depending on the gene it is native to and the regulatory proteins that bind the G4-forming region. G4s have been identified in RNA as well. Such structures form more readily than DNA G4s, as they do not need to compete with duplex formation. In 2010, telomeric repeat-containing RNA (TERRAs) were characterized to form RNA-G4s within living human cells (Xu et al. 2010). RNA-G4s have been described in vivo, and studies found that such structures are localized at the ends of chromosomes suggesting that RNA-G4s occur in telomeric regions of the human

genome. A few years later, a G4 structure-specific antibody, BG4, was engineered to investigate G4 existence in human cells and within individual chromosomes (Biffi et al. 2013). G4s were visualized and suggested to form both within and outside the telomeric regions of chromosomes, which further validates the presence of promoter G4s. These breakthroughs indicate that G4s possess biological roles in replication and in regulating gene expression both transcriptionally and translationally.

Extensive evidence describes the folding of non-B-DNA arrangements in vitro throughout the genome. Following the completion of Human Genome Project in 2003, the understanding of G4s existing throughout the genome has increased exponentially. Computational analyses of putative G4 formation using the consensus sequence, $(G_3N_{1-7})_3G_3$, displayed a high density exists within promoters and 5' untranslated regions (UTRs) (Verma et al. 2008). This was examined within the genomes of four species: human, chimpanzee, rat, and mouse, which suggests that the putative G4s described are evolutionarily conserved. In 2015, genome-wide sequence analysis was performed within the human genome from B lymphocytes (Chambers et al. 2015), and over 700,000 distinct G4s were found, about 450,000 of them were not previously accounted from by computational methods using the putative G4 consensus sequence. The distinct G4s characterized were highly associated with oncogenes and tumor suppressor genes and were found in high densities regulatory regions.

It is estimated that approximately 40% of human gene promoters contain putative G4-forming regions (Huppert & Balasubramanian 2006). Localization has been reported to be significantly higher in proto-oncogenes than in housekeeping or tumor suppressor genes (Eddy 2006). Gene promoters tend to be rich in guanines where chromatin

structure is open, allowing for easier transcriptional activation. G4 formation is permitted in promoter regions when DNA transcriptional machinery (RNA Pol II and transcription factors) bind to double-stranded DNA (dsDNA). Strand separation occurs due to the negative supercoiling that occurs behind the transcriptional machinery, upstream from the transcriptional start site (TSS) (Wang & Dröge 1996; Kramer & Sinden 1997; Champoux 2001; Brooks & Hurley 2009). Torsional stress is usually relieved by topoisomerases. However, competition arises with other regulatory binding proteins which may preclude access for topoisomerase binding (Champoux 2001). Also, this form of regulation has been reported as insufficient in removing all supercoiling that is generated during transcription which would allow secondary structure formation (Wang & Dröge 1996).

As mentioned, there exists a plethora of G4s within oncogenes and tumor suppressor genes that have previously been characterized such as MYC, kRAS, hRAS, nRAS, c-KIT, Bcl-2, pRB, VEGF, HIF-1, hTERT, and more. In fact, there is at least one example of a G4 forming and acting upon critical genes responsible for each of the ten hallmarks of cancer (Figure 1-2) (Brooks & Hurley 2010; Cogoi & Xodo 2006; Morgan & Brooks 2016; Morgan et al. 2016; Onyshchenko et al. 2009; Sun et al. 2014; Phan et al. 2004; Huang et al. 2012; Bonnal et al. 2003; Ambrus et al. 2006; Agarwala et al. 2013; Williamson 1994). Most promoter G4s have been found to act as transcriptional silencers through the disruption of regulatory protein binding. The most characterized G4 to date is within the MYC promoter, which is responsible for downregulating transcription via obstructing Sp1, CNBP, and hnRNP-k protein binding (Brooks & Hurley 2009). In contrast, there exists a G4 that forms within the Bcl-2 promoter that displaces transcriptional suppressor, WT1, binding resulting in activation instead (Brooks et al.

2010). Therefore, structure stabilization for transcriptional control has become a key interest for G4 research and drug discovery programs utilizing a small molecule approach.

G4 research heavily concentrates on cancer drug discovery. With MYC being the first promoter G4 characterized and the most extensively studied structure to date, there have been numerous attempts to target the structure with a large variety of small molecules. These studies led to the design of a first-in-class G4-interacting drug, Quarfloxin (Drygin et al. 2009). This compound bound MYC G4 with higher affinity than universal G4-binding agents and led to decreased mRNA expression isolated from a mouse xenograft model of human colorectal cancer cells. Unfortunately, due to high plasma protein binding, Quarfloxin did not proceed past Phase II trials. To date, Quarfloxin and compound CX-5461 are the only G4-stabilizing agents to make it to clinical trials with CX-5461 being the most recent and is currently being tested (Xu et al. 2017). However, prospective drug discovery designs utilizing G4s to control gene expression through transcription have potential. The current pitfalls are due to the gap in knowledge of bridging physiological G4 formation with that in small molecule screening which takes place outside of cells. We aim to gain such knowledge through the following aims:

Specific Aim 1. Discern biologically relevant G4s by examining the role of epigenetic modifications on G4 stability and recognition, utilizing physiological conditions of a cell's nucleus extracellularly, and distinguish G4 biological activity on transcription.

Specific Aim 2. Screen for small G4-stabilizing agents for anti-cancer or neuroprotective agents using physiological conditions extracellularly.

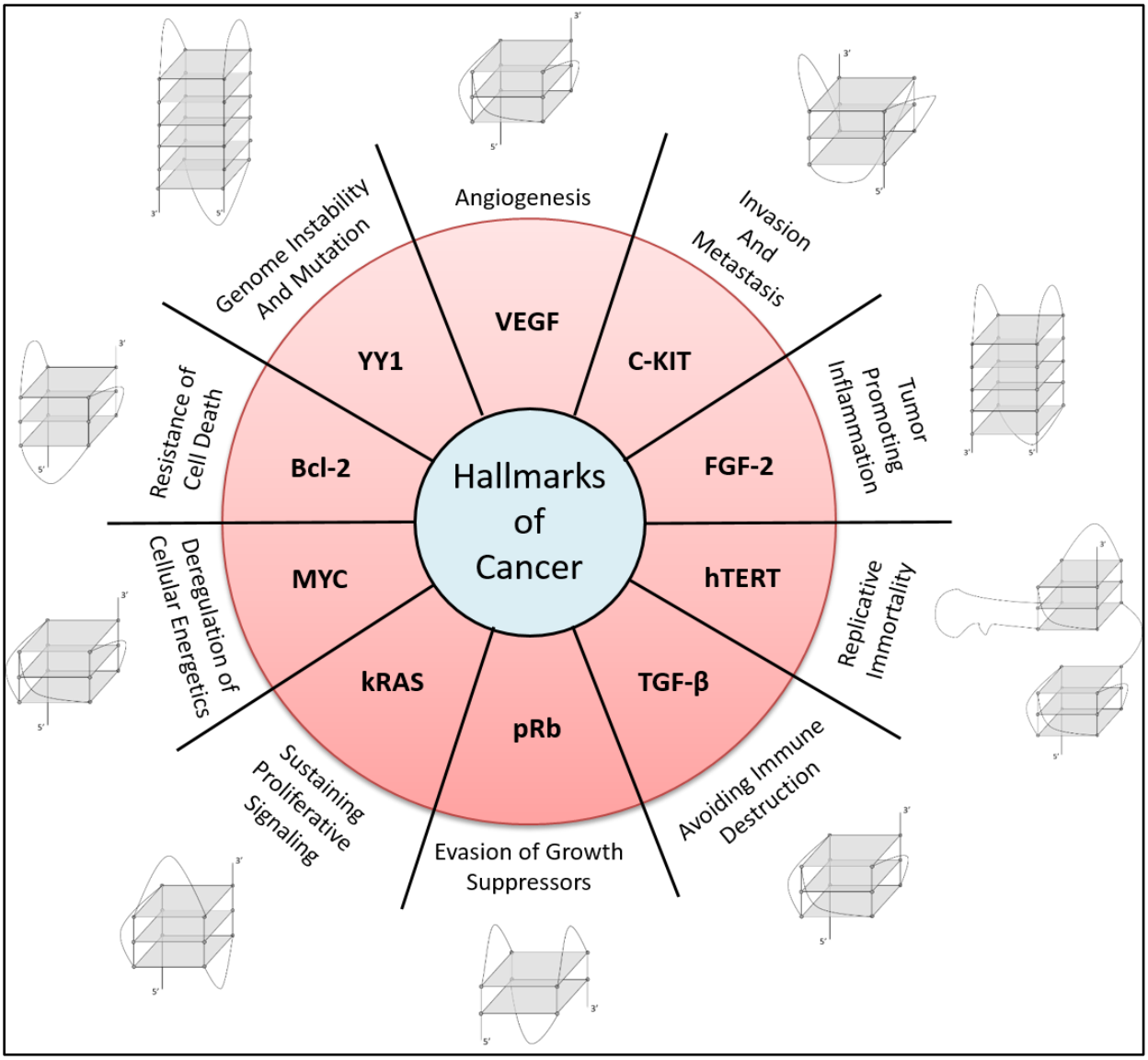


Figure 1-2. G4s in the hallmarks of cancer.

CHAPTER 2. MATERIALS & METHODS

2.1. Reagents

Most oligonucleotides were synthesized and purchased from Eurofins MWG Operon, LLC (Louisville, KY) with the exception of 5-hmC oligonucleotides which were purchased from Midland Certified Reagent Co., Inc. (Midland, TX) (Table 2-1). Acrylamide/bisacrylamide (29:1) solution and ammonium persulfate were purchased from Bio-Rad laboratories (Hercules, CA), and N,N,N',N'-tetramethylethylenediamine was purchased through Thermo Fisher Scientific (Pittsburgh, PA). Unless specified, all other chemicals were purchased from Sigma-Aldrich (St. Louis, MO).

2.2. DNA Solvation & Concentration Determination

Upon receiving, DNA was resuspended in water and thoroughly vortexed for 2 min before being heated at 95 °C for 10 min. Concentrations were determined, while the DNA was still hot, using the ratio of the observed A260 via a NanoDrop2000 to the E260 of the oligonucleotide (as reported by IDT Oligo Analyzer).

Table 2-1. List of all oligonucleotide sequences.

$VEGF_{KO}$	5'CCGTTGCGTGCCGTTTGCGTTGTC3'
$VEGF_{1-4}$	5'CCGGGGCGGGCCGGGGGCGGGGTC3'
$VEGF_{mod1}$	5'CCGGGG(5-HM-dC)GGGCCGGGGGCGGGGTC3'
$VEGF_{mod2}$	5'CCGGGGCGGG(5-HM-dC)CGGGGGCGGGGTC3'
$VEGF_{mod3}$	5'CCGGGGCGGGCCGGGGG(5-HM-dC)GGGGTC3'
MYC_{KO}	5'GCGCTTATGTTGAGTGTGTTGAGTGTGTTGAAGGTGTTGAGGAGAC3'
MYC_{KO} EMSA	5'[6~FAM]GCGCTTATGTTGAGTGTGTTGAGTGTGTTGAAGGTGTTGAGGAGAC3'
MYC	5'GCGCTTATGGGGAGGGTGGGGAGGGTGGGGAAGGTGGGGAGGAGAC3'
MYC EMSA	5'[6~FAM]GCGCTTATGGGGAGGGTGGGGAGGGTGGGGAAGGTGGGGAGGAGAC3'
MYC FRET	5'[6~FAM]GCGCTTATGGGGAGGGTGGGGAGGGTGGGGAAGGTGGGGAGGAGAC[TAMRA~6~FAM]3'
$G4_{near\text{-}KO}$	5'AGGGCTTTTTGGGAAGAGTGAAGAGGGGGAGG3'
$G4_{near\text{-}KO}$ EMSA	5'[6~FAM]AGGGCTTTTTGGGAAGAGTGAAGAGGGGGAGG3'
$G4_{near}$	5'AGGGCGGTGTGGGAAGAGGGAAGAGGGGGAGG3'
$G4_{near}$ EMSA	5'[6~FAM]AGGGCGGTGTGGGAAGAGGGAAGAGGGGGAGG3
$G4_{mid\text{-}KO}$	5'CGGGGAGAAGGAGTTTGCCGGGCCGGGCCGGCGGGGGAGGAGCGTTTGCCGGGC3'
$G4_{mid\text{-}KO}$ EMSA	5'[6~FAM]CGGGGAGAAGGAGTTTGCCGGGCCGGGCCGGCGGGGGAGGAGCGTTTGCCGGGC3'
$G4_{mid}$	5'CGGGGAGAAGGAGGGGGCCGGGCCGGGCCGGCGGGGGAGGAGCGGGGGCCGGGC3'
$G4_{mid}$ EMSA	5'[6~FAM]CGGGGAGAAGGAGGGGGCCGGGCCGGGCCGGCGGGGGAGGAGCGGGGGCCGGGC3'
$G4_{mid}$ DMS	5'[6~FAM]TTTTTTTCGGGGAGAAGGAGGGGGCCGGGCCGGGCCGGCGGGGGAGGAGCGGGGGCCGGGCTTTTTTT3'
$G4_{mid}$ FRET	5'[6~FAM]GCGGGGAGAAGGAGGGGGCCGGGCCGGGCCGGCGGGGGAGGAGCGGGGGCCGGGC[TAMRA~6~FAM]3'
$G4_{mid}$ DMS Forward Primer	5'-[6~FAM]GATGCGTTCCGCGCTCGA-3'
$G4_{mid}$ DMS Reverse Primer	5'-[6~FAM]AGTCCCTCCTCCCGCCAA-3'
Mut A	5'CGTTGAGAAGGAGGGGGCCGGGCCGGGCCGGCGGGGGAGGAGCGGGGGCCGGGC3'
Mut B	5'CGGGGAGAAGGAGGTGGCCGGGCCGGGCCGGCGGGGGAGGAGCGGGGGCCGGGC3'
Mut C	5'CGGGGAGAAGGAGGGGGCCGTGCCGGGCCGGCGGGGGAGGAGCGGGGGCCGGGC3'
Mut D	5'CGGGGAGAAGGAGGGGGCCGGGCCGTGCCGGCGGGGGAGGAGCGGGGGCCGGGC3'
Mut E	5'CGGGGAGAAGGAGGGGGCCGGGCCGGGCCGGCGGTGGAGGAGCGGGGGCCGGGC3'
Mut F	5'CGGGGAGAAGGAGGGGGCCGGGCCGGGCCGGCGGGGGAGGAGCGGTGGCCGGGC3'
Mut G	5'CGGGGAGAAGGAGGGGGCCGGGCCGGGCCGGCGGGGGAGGAGCGGGGGCCGTGC3'
$G4_{far\text{-}KO}$	5'AAGGTGTGGCTGTTGCGGTCTAGTGTGGCGAGCCGTGCC3'
$G4_{far\text{-}KO}$ EMSA	5'[6~FAM]AAGGTGTGGCTGTTGCGGTCTAGTGTGGCGAGCCGTGCC3'
$G4_{far}$	5'AGGGGTGGCTGGGGCGGTCTAGGGTGGCGAGCCGGGC3'
$G4_{far}$ EMSA	5'[6~FAM]AGGGGTGGCTGGGGCGGTCTAGGGTGGCGAGCCGGGC3'
$G4_0$	5'CGGGCGGGAGCGCGGCGGGCGGGCGGGC3'
$G4_1$	5'AGGGGCGGGCGCGGGAGGAAGGGGGCGGGAGCGGGGC3'
$G4_1$ FRET	5'[6~FAM]AGGGGCGGGCGCGGGAGGAAGGGGGCGGGAGCGGGGC[TAMRA~6~FAM]
$G4_2$	5'CGGGCCAGGGAGCGGGGCGGAGGGGGCGGTCGGGT3'
$G4_3$	5'GCTGGGGTCCGCGACGGGGTGGGGGCTCCCGGGGAAC3'

2.3. Electronic Circular Dichroism (ECD)

The G4 oligonucleotides (5 µM) were diluted in 50 mM Tris-HCl (pH 7.4) with varying amounts of KCl and co-solvents (0 - 40%). The co-solvents utilized were (1) osmolytes: acetonitrile (MeCN), glucose, sucrose, dextran sulfate, and ficoll and (2) molecular crowding agents: polyethylene glycol 300 (PEG), glycerol, and extracted nucleoplasm. Samples were heated at 95 °C and snap cooled on ice for 5 min each and repeated up to five times depending on the DNA. Concentration was determined using the ratio of A260 observed with a NanoDrop2000 by the E260 of the oligonucleotide. Spectra were collected with an Olis DSM-20 spectropolarimeter equipped with a ECD 250 Peltier cell holder (Bogart, GA). Recordings were made over the wavelength range 225-350 nm (every nm) and at increasing temperatures (20-100 °C, every 10 °C, with a 1 min. hold at temperature before spectra were recorded) in 1 mm quartz cuvette. The ordinate is reported as molar ellipticity (converted from millidegrees based on sample concentration); T_M's were determined by performing a singular value decomposition (SVD) analysis available with the Olis GlobalWorks software, followed by non-linear regression fitting sigmoidal inhibition dose-response on GraphPad Prism software (La Jolla, CA) in order to determine the 50% melting point (T_M).

2.4. Electrophoretic Mobility Shift Assay (EMSA)

FAM-labeled oligonucleotides, MYC and $kRAS_{mid}$, were prepared in 10 mM KCl or 100 mM KCl, 50 mM Tris-HCl (pH 7.4), and one of the following: 2-100% extracted nucleoplasm, 20% sucrose, 20% MeCN, 10% dextran sulfate, 20% PEG, or 10 µM G4-stabilizing compounds. The solutions were each denatured by heating to 95 °C for 5 min, and rapidly cooled on ice for 5 min, and repeated for one to five cycles, depending on the

G4, to induce structure formation. Upon addition of non-denaturing loading dye (1:1 ratio of 6x loading dye and 50% glycerol), the samples were loaded onto a 10-16% native polyacrylamide gel (10-16% of acrylamide-bisacrylamide solution, 1x TBE buffer, 10 mM KCl, 1% APS, 0.1% TEMED). After running at 100 V, the gel was visualized under blue light LED using a FOTO/Analyst® Investigator FX Imager (Hartland, WI).

2.5. Cell Culture

Burkitt's lymphoma cells, RAJI and CA46, were maintained in Roswell Park Memorial Institute (RPMI) 1640 medium. Pancreatic cancer, Panc-1, MiaPaCa-2, and AsPc-1, and human embryonic kidney, HEK-293 cells were preserved in Dulbecco's modified eagle's medium (DMEM). Neuroblastoma cells, SH-SY5Y, were kept in Dulbecco's modified eagle/nutrient mixture F-12 medium (DMEM/F-12). All media types were supplemented with 10% fetal bovine serum and 1x penicillin/streptomycin solution, and cells were incubated at 37 °C, in a humidified environment containing 5% CO_2. Each cell line was obtained from ATCC (Manassas, VA).

2.6. Cellular Nucleoplasm Extraction

MiaPaCa-2 cells were grown to ~90% confluency in 75 mL flasks, trypsinized, and collected in a 50 mL conical tube. Cells were resuspended and centrifuged at 4,000 rpm for 5 min. in ~5 packed cell volume (pcv) hypotonic buffer solution (10 mM HEPES, pH 7.9, 1.5 mM $MgCl_2$, 10 mM KCl) after adding 0.2 mM phenylmethylsulfonyl fluoride (PMSF) and 0.5 M dithiothreitol (DTT). The pellet was then resuspended in 3 pcv hypotonic buffer solution, and the cells were allowed to swell on ice for 10 min. Cells were homogenized and spun at 4,000 rpm for 15 min. and were then resuspended in ½ packed nuclear volume (pnv) low-salt buffer solution (20 mM HEPES, pH 7.9, 25% glycerol, 1.5

mM $MgCl_2$, and 0.2 mM EDTA) after adding 0.2 mM PMSF and 0.5 M DTT. Slowly, ½ pnv high-salt buffer solution (add 0.8 M KCl to low-salt buffer stock) was added in a drop-wise fashion upon addition of 0.2 mM PMSF and 0.5 M DTT. Cytoplasmic extract was incubated at room temperature for 30 min with continuous mixing and then centrifuged at 14,000 rpm for 30 min; supernatant was collected as the nuclear extract.

2.7. Dimethyl Sulfate (DMS) Footprinting

FAM-labeled oligonucleotides (2 μM) were prepared in 50 mM Tris-HCl (pH 7.4), 10 or 100 mM KCl, alone or with one of the following: 2-5% extracted nucleoplasm, 20% sucrose, 20% MeCN, 10% dextran sulfate, 20% PEG or 10 μM G4-stabilizing compounds. Samples were denatured at 95 °C for 5 min and then cooled at 4 °C for 5 min. This was repeated for one to five cycles, depending on the G4, to allow structure induction. Oligonucleotide solutions were then treated with 1% DMS (one part DMS to 4 parts ethanol) and incubated at room temperature for 15 min. The reaction was stopped with 5% β-mercaptoethanol. The samples were separated on a 10% native gel at 100 V. Bands were visualized on a FotoDyne Imager, excised with scalpel, and DNA was eluted from the gel overnight. Upon filtration by centrifugation, the DNA underwent ethanol precipitation (70% EtOH, 0.3M NaOAc) overnight at -20 °C. Solutions were centrifuged at 14,000 rpm for 30 min and resuspended in 10% piperidine. This treatment was incubated at 90 °C for 30 min and ethanol precipitation overnight as previously described. Samples were spun at 14,000 rpm for 30 min, dried, and then resuspended in 50 μL DMS dye (95% formamide, 0.01% bromophenol blue, 20 mM EDTA). Finally, solutions were heated at 95 °C for 5 min and placed on ice until being loaded onto a 16% denaturing

polyacrylamide gel run at 500 V. The gel was then visualized on a FOTO/Analyst® Investigator FX Imager (Fotodyne, Inc., Hartland, WI).

2.8. Fluorescence Resonance Energy Transfer (FRET) Melt

A dual-labeled DNA oligomer probe (200 nM) bearing the G4-forming regions of the MYC, $kRAS_{mid}$, or Bcl-2 promoters were used to screen compounds of interest (10 μM). DNA was diluted in 10 mM Tris-HCl + 90 mM LiCl, 10 mM sodium cacodylate, heated at 95 °C and rapidly cooled on ice. Annealed G4 DNA was added to a 96-well PCR plate with or without test compounds and incubated at room temperature for 30 min. Fluorescence was recorded from 25-95 °C, at every degree after a 30 sec hold on a Bio-Rad CFX Connect real-time PCR machine (Hercules, CA). Over 1,600 compounds were screened from NCI Diversity Set III for stabilizing kRAS G4mid, and all other compounds screened were from an “in-house” 96-well plate containing known pan-G4-stabilizing agents, as well as, compounds isolated or synthesized at the University of Mississippi.

2.9. MTS Cell Viability

All cells were plated in 96-well plates in 90 μL of media. Burkitt’s lymphoma cells, RAJI and CA46, were plated at 75,000 cells/well, pancreatic cancer cells, Panc-1, MiaPaCa-2, and AsPc-1, were plated at 10,000 cells/well and were allowed to attach for 24 hr, neuroblastoma cells, SH-SY5Y, were plated at 60,000 cells/well and allowed to attach for 48 hr. Following the attachment time, compounds (10 μL of 10x) of interest were added over a 5-6 log range and incubated at 37 °C with 5% CO_2 for 24-72 hr depending on the range listed above. Changes in cellular viability were determined by the MTS assay: 2 mL of sterile filtered 2 mg/mL MTS in PBS with 5% PMS added immediately before use, plates were then incubated at 37 °C with 5% CO_2 for 2-4 hr before recording absorbance

using a BioTek Synergy 2 spectrophotometer (BioTek Instruments, Inc., Winooski, VT). The recorded absorbance of the formazan dye at 490 nm were normalized to control cells; GraphPad Prism non-linear regression fitting sigmoidal inhibition dose-response was used to calculate the IC_{50}.

2.10. Quantitative Real-Time PCR (qRT-PCR)

Pancreatic cancer cells (BxPc3, Panc-1, MiaPaCa-2, AsPc-1) were plated at 100,000 cells/well in a 12-well plate and incubated overnight at 37 °C 5% CO_2. DMSO or compounds of interest were added to the seeding media final concentrations approximating the IC_{50} (as determined by MTS cell viability assay). After 24 hr, samples were resuspended in lysis buffer (guanidine thiocyanate mixed with EtOH, and 20 µL 14.3 M β-mercaptoethanol added to each 1000 µL buffer used), and RNA was extracted using Thermo Scientific GeneJET RNA purification kit. RNA concentrations were determined using a NanoDrop 2000 (Thermo Scientific), and only samples with 260/230 values >2 were used to synthesize cDNA. 200ng of cDNA was synthesized using iScript™ cDNA synthesis kit (Bio-Rad), and qRT-PCR was run on an Applied Biosystems 7500 real-time PCR detection system (ThermoFisher Scientific) using TaqMan primers from ABI (kRAS: Hs00364284_g1, MYC Exon 2: Hs00153408_m1, MYC Exon 1: Hs01562521_m1, Bcl-2: Hs00608023, GAPDH: Hs02758991_g1). mRNA expression was normalized to GAPDH, and again to vehicle controls using the $\Delta\Delta C_q$ method. Experiments were run in biological duplicates and technical triplicates; statistical significance was determined by one-way ANOVA with post-hoc Tukey analysis.

2.11. Plasmid DMS Footprinting

DMS footprinting was performed in a cell-free environment with the supercoiled kRAS plasmid having the promoter region from +50 to -324 described above. The plasmid (500 ng) was combined with 100 mM KCl alone or with 100 μM TMPyP4 and incubated at 37 °C overnight, after which, DMS and piperidine treatment will be performed as described above. 1.25 units HotStart polymerase, 1x standard Taq reaction buffer, 200 μM dNTPs, and 200 nM FAM-labeled primers (Forward: 5'-GATGCGTTCCGCGCTCGA-3', Reverse: 5'-AGTCCCTCCTCCCGCCAA-3') will be added for PCR amplification (Initial denaturation: 95 °C for 30 sec; 30 cycles: 95 °C 20 sec, 58 °C 30 sec, 68 °C 1 min; Final extension: 68 °C 5 min). The products were separated on a denaturing PAGE gel and visualized on a FOTO/Analyst® Investigator FX Imager (Fotodyne, Inc., Hartland, WI).

2.12. Plasmid Construction

Using the pGL4.17 plasmid (Promega) as a backbone, several luciferase plasmids were constructed. Promoter-driven, SV40, MYC Del4, and Bcl-2 plasmids were purchased through Addgene (Cambridge, MA). The kRAS promoter regions of interest were inserted between the Nhe I and HIND III cut sites. $G4_{near}$, $G4_{FL}$, and $G4_{mid}$ were constructed by Operon while we synthesized the "no G4", mut near and mut $G4_{mid}$ constructs. Using the primers, 5'-GACGCTAGCTCAGCCGCTCCCTCTCGTAC-3' and 5'-GACAAGCTTACCTTCGCCGCCGCCACTGC-3', the "no G4" plasmid was constructed by PCR from the $G4_{near}$ plasmid. The PCR product was ran on a 1.5% agarose gel, extracted, and cut by Nhe I and HIND III enzymes overnight. Then, the product was ligated into an EV plasmid. The mut $G4_{near}$ and mut $G4_{mid}$ plasmids were constructed using $G4_{FL}$ and the Q5® Site-Directed Mutagenesis kit (New England BioLabs, Inc.). Mut

$G4_{near}$ required primers 5'-TTTGGGAAGAGGGAAGAGGGG-3' and 5'-AAGCCCTCAGCCGCTCCCTCTC-3', and the mut $G4_{mid}$ used primers 5'-GCGGGGGAGGAGCCCGGCGGAGGAAG-3' and 5'-CGGCCCGGCCCGGCTCCCCGCCGGCG-3'. Product size was validated using an agarose gel.

2.13. Transfection and Luciferase Assays

HEK-293 cells were plated 200,000 cells/well in 6-well plates and allowed to attach overnight. SH-SY5Y cells were plated 100,000 cells/well in 24-well plates and allowed to attach for 48 hr. Each well was then co-transfected with the plasmid of interest (200 ng kRAS plasmids, 450 ng Del4 and Bcl-2) and the reference plasmid pRL-SV40, in equal amounts to plasmid of interest, formed into micelles with FuGene HD (Promega) in a 3:1 ratio with cells in OptiMEM medium (Invitrogen; Grand Island, NY). An unmodified pGL4.17 plasmid (Empty Vector, EV) was used as a negative control vector, at equal amounts to plasmid of interest. A promoter-driven plasmid, SV40, was used as a positive control at equal amounts to plasmid of interest. Transfections were maintained at 37 °C in 5% CO_2 overnight before media was replaced with DMEM alone or with test compound. After 24 or 48 hrs of incubation, the expression of firefly, with respect to Renilla, luciferase was established using the Dual-Reporter Luciferase assay kit (Promega); luminescence was measured with a Lumat LB9507 luminometer (Berthold Technologies). All experiments were performed in biological duplicate with technical duplicates. Statistical significance was determined by a one-way ANOVA with post-hoc Tukey analysis or a two-way ANOVA with Bonferroni post-test analysis (in SH-SY5Y samples).

2.14. Neuroprotection Studies

Neuroblastoma cells were plated at 90,000 cells/well in 96-well plates and allowed to attach for 48 hr. For serum starvation, our environmental stressor, media was removed and replaced with new media containing 1% or 10% of fetal bovine serum (FBS); inflammatory stress was emulated upon 250 µM H_2O_2 incubation 3 hr prior to MTS addition. Neurotoxicity was induced with 100 µM 6-hydroxydopamine (6-OHDA). Compounds of interest were added to each condition listed in the same way as in the MTS cell viability assay described above. Neuroprotection was measured by observing the changes in cell viability. All experiments were performed in biological duplicate with internal technical triplicates.

CHAPTER 3. PHYSIOLOGICAL MIMICRY

3.1. Introduction

G4s are common throughout the genome with high abundances in promoters and 5'UTRs in oncogenes and tumor suppressor genes (Verma et al. 2008; Chambers et al. 2015). G4s are highly polymorphic, both across regions of DNA and within one sequence, and vary in the number of stacked tetrads, as well as in loop directionality. It is suggested that short loop lengths are favorable, but there are G4s having been characterized harboring long loops like kRAS, hTERT, and Bcl-2. These non-canonical DNA structures have been characterized as transcriptional regulators, as either activators or suppressors depending on the gene involved. Most G4s identified are transcriptional suppressors by hindering RNA polymerase activity and/or transcription factor binding. Such structures span the hallmarks of cancer, thus, most research focuses on utilizing small G4-stabilizing molecules in order to decrease tumor activity.

During replication or transcription, the complimentary strands of gene promoter regions are separated. Separation of the dsDNA helix to single stranded DNA (ssDNA) has the potential to form G4 structures (Ahmed et al. 1994; Michelotti, Gregory A. 1996; Rangan et al. 2001; Simonsson et al. 2000). At physiological pH, G4s have previously been described to form complexes with crowding reagents aiding in the stability of structure formation. Also, these studies have confirmed that G4s are highly polymorphic and differ between genes. The average topology of G4s vary depending on the

environment as well as the sequences, specifically pH, dehydrating conditions, and epigenetics (Miller et al. 2010; Zhang et al. 2010; Dailey et al. 2010).

As the interest in G4 function in biological systems grows, understanding the structure thermodynamics in the presence of biomolecules will help to clarify what occurs in physiological conditions. Biomolecules possess functions within living cells and inhabit a marked portion of a cell's volume (approximately 20-40%) (Miyoshi et al. 2006). It has been reported that molecular crowding affects the stability of DNA as a result of changes in hydration (Goobes et al. 2003; Spink & Chaires 1999). In particular, PEG has been found to induce alterations in G4 structure with K+ present. This was due to fewer hydroxyl groups within the co-solvent as these groups affect DNA hydration. Molecular crowding with a K+ cation present causes dehydration and stabilization of G4s while it destabilizes duplex DNA as Watson-Crick base pairs consist of more hydration sites than Hoogsteen-base pairs (Misra & Draper 1998). This suggests that non-canonical DNA structures can be favored over B-DNA duplexes depending on the sequence and surrounding nuclear conditions which directs structural polymorphisms found to be regulated via hydration states. However, previous studies have not explored the effect epigenetic DNA modifications has on G4 structures in the presence of the most relevant co-solutes to recapitulate intranuclear higher order formations.

Relevant in the study of genetics, epigenetics is the study of cellular traits that are influenced by external or environmental factors. The genes producing the cellular traits can be upregulated or downregulated based on the effects the influential factors have on transcription (Simmons 2008). Not to be confused with traditional studies in genetics, which focuses on base sequence changes, the central focus of epigenetics is on

modifications to a DNA sequence not involving base sequence changes (Carey 2011). Modifications such as methylation or hydroxymethylation to bases in a DNA sequence have been discovered as epigenetic external factors regulating gene expression. Enzymes of the methyltransferase and ten-eleven-translocation (TET) family of oxygenases typically generate the methylated or hydroxymethylated bases, respectively (Bhavsar-Jog, Van Dornshuld, Brooks, et al. 2014). Unique regions in a genome prone to these types of modifications are CpG islands. CpG islands consist of repeats of a cytosine being placed next to a guanine in the same sequence with only a phosphate separating the two nucleotides (Josse et al. 961; Morrison, J. M. 1967). The occurrence of CpG islands in mammals is quite low, less than 1% in humans (Scarano, E. 1967). Low occurrence of CpG islands can be linked to the likelihood of cytosines being methylated causing deamination to thymine. However, when CpG islands occur, the chance of external modification on a cytosine increases the probability of altering gene expression (Jabbari & Bernardi 2004). Hypermethylation can suppress gene transcription important for cellular growth and function such as tumor suppressor genes while hypomethylation can upregulate oncogenes promoting tumorigenesis. These actions can be especially important in cancer-causing genes containing G4-forming regions as these regions exist within CpG islands and are highly susceptible to epigenetic modifications, specifically 5-hydroxymethylated cytosines (5-hmC) (Halder et al. 2010).

Non-canonical DNA structures have been discovered to exist within many oncogene and tumor suppressor gene promoters. Previous ones that have been characterized include MYC, Bcl-2, hRAS, pRb, c-KIT, PDGF-A, and VEGF genes. Of these known G4s, MYC and VEGF have been extensively studied. The G4-forming region

within the VEGF promoter occurs within a CpG island. This is not true for the MYC promoter, thus VEGF is an ideal candidate for investigating the role of epigenetic modifications on G4 stability in the presence of co-solutes. On the other hand, MYC is an ideal candidate for exploring the incorporation of co-solvents into G4 drug discovery screens in order to simulate intranuclear conditions in vitro.

Several compound classes have been identified to bind and stabilize G4s. Many molecules universally bind G4 structures and impact gene expression in vitro with little influence in cells or in vivo. It is challenging to classify new chemical scaffolds as selective G4 stabilizing agents; the techniques utilized in order to ascertain small molecule-mediated stabilization of G4s are insufficient and not optimal as they are performed under simple conditions. Standard initial protocols for screening G4-stabilizing compounds, such as FRET melt and ECD, are operated in buffer in vitro with ssDNA and the addition of a salt (KCl or NaCl). This allows for structure formation, but it does not fully coincide with physiologically relevant conditions or the biologically active G4 isoforms. Thus, drug discovery and development programs are hindered, as there have been a lack of lead compounds that selectively modulate G4-driven genes. To date, one agent triumphed to clinical trials for stabilizing the G4 associated with transcription factor MYC which plays a role in tumorigenesis and disease progression (Brooks & Hurley 2010). However, Quarfloxin (CX-3543) did not succeed phase II trials due to high albumin binding (Balasubramanian & Hurley 2011). Consequently, modern G4 drug discovery efforts require attention and optimization to ensure in vitro success that will correlate to increased favorability in vivo.

Applying G4 stabilization to gain transcriptional control of gene regulation in order to suppress oncogene expression is a promising cancer therapeutic approach. Stabilization of these structures within promoter regions act as blockades to RNA polymerase and other gene regulatory proteins. Enhancing the formation and stability is therefore potentially therapeutic, and as these structures are more “globular” than traditional dsDNA, future drug discovery programs can focus on particular genes/structures with increased selectivity as compared to dsDNA. Extracellularly, stabilization can occur by the addition of alkali cations, modifying water content, or by integrating molecular crowding. Intracellularly, stabilization can be attained pharmacologically by applying small molecules. Ligand-mediated G4 stabilization has been a growing drug discovery program since their initial characterization. Gaining more knowledge on how these structures form will aid in the development of future drug discovery efforts. Also, pharmacological modulation of G4-driven genes with small molecules will have higher success (decreased false positives and false negatives) when transitioning from in vitro to in vivo upon utilization of co-solvents to simulate nuclear conditions of a cell. Dehydrating and molecular crowding agents, specifically MeCN and PEG, have been studied previously and were found to increase G4 thermodynamic stability. Modifying water content has also been shown to enhance binding affinity of ligands to DNA (Winn et al. 2009). Hence, the addition of co-solvents to G4 screening protocols would enable anhydrous molecular crowding and simple dehydration, mimicking intracellular conditions, and result in a minimization of false positives for ligand binding.

We hypothesize that 5-hmC-containing G4s may alter G4 stability and recognition by G4-stabilizing small molecules and regulatory binding proteins. We determined this by examining the role of cationic strength, pH, and water content on structures containing 5-hmC modifications in G4 loops. We sought to define whether 5-hmC modifications could influence the average topology and stability of VEGF G4 structures which contains natural CpG islands. Also, we investigated the enhancement of current G4 drug discovery programs through the use of co-solvents to mimic physiological conditions of a cell's nucleus. Here, we hypothesize that the use of co-solvents to recapitulate physiological conditions extracellularly will increase success in G4 drug discovery programs when testing compounds outside of cells, confirming the mechanism inside of cells, as well as in examining in vivo efficacy. Collectively, these works can be divided into the following aims:

Specific Aim 1. Quantify the thermodynamic properties and underlying equilibrating G4 structures containing 5-hmC modifications.

Specific Aim 2. Determine the role of co-solvents on G4 folding, stability, and polymorphic nature and apply this physiological mimicry to G4 drug discovery programs.

3.2. 5-hydroxymethylcytosine Modification Effects on G4 Formation and Stability

Vascular endothelial growth factor (VEGF) produces a protein that stimulates angiogenesis allowing oxygen restoration to tissue (Senger, 1983). When overexpressed, VEGF is responsible for the rapid development of new blood vessels to cells, specifically tumor cells. This can cause vascular disease and can help tumors metastasize. VEGF is a key player in tumorigenesis for many cancer cell types making it a promising target for drug therapy. The promoter within this oncogene contains a main cis-element located -85 to -50 bp from the TSS. This element is a G4 forming region consisting of five contiguous runs of guanines. Previous studies have described the G4 structures that form with the first four runs in ssDNA. Besides enabling G4 formation, this G/C rich region close to the TSS contains a CpG island favorable for cytosine methylation or hydroxymethylation. Thus, we utilized VEGF oligonucleotides containing 5-hmC modifications, compared to wild-type ($VEGF_{1-4}$), on three different cytosines with the potential for natural cytosine modification, each involved in loop formation within the G4 (Table 3-1) (Figure 3-1A). The ssDNA oligonucleotides containing modifications were predicted based on CpG mono- or di-nucleotide motifs. It is important to note that the cytosine modified for $VEGF_{mod2}$ was chosen as cytosine modifications occur on the most distal nucleotide in a $CCG_{(1-2)}$ motif (Stroud et al. 2011).

Table 3-1. VEGF modification oligonucleotide sequences.

$VEGF_{KO}$	5'CCGTTGCGTGCCGTTTGCGTTGTC3'
$VEGF_{1-4}$	5'CCGGGGCGGGCCGGGGGCGGGGTC3'
$VEGF_{mod1}$	5'CCGGGG(5-HM-dC)GGGCCGGGGGCGGGGTC3'
$VEGF_{mod2}$	5'CCGGGGCGGG(5-HM-dC)CGGGGGCGGGGTC3'
$VEGF_{mod3}$	5'CCGGGGCGGGCCGGGGG(5-HM-dC)GGGGTC3'

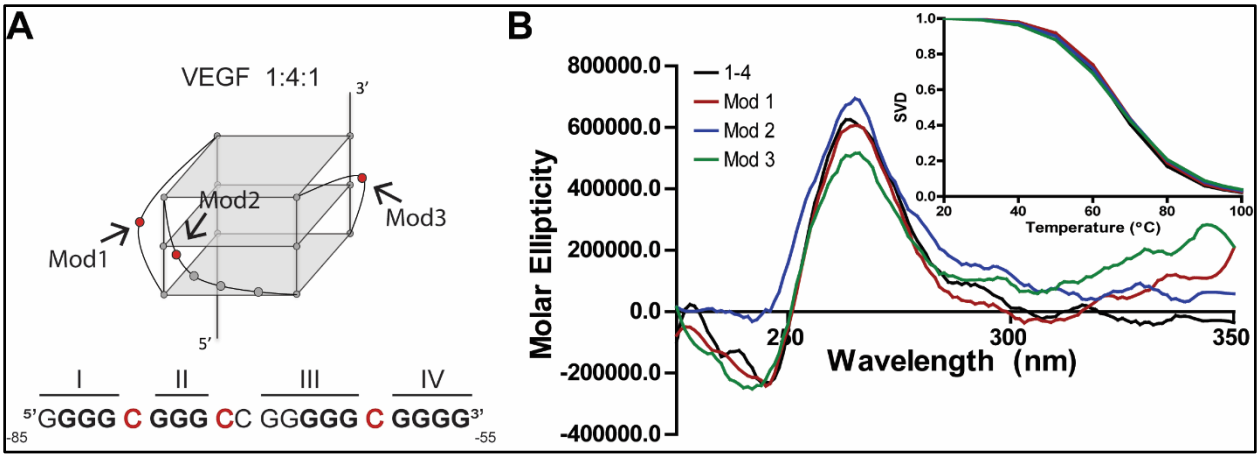


Figure 3-1. 5-hmC modifications within VEGF G4 allow stable G4 formation. Wild-type VEGF forms a parallel G4 with cytosines contained within the loops that may possess 5-hmC modifications. These potential sites have each been modified in different oligonucleotides named mod 1, 2, and 3 depending on the location of the 5-hmC (A) ECD spectra shows stable G4 formation with each of the modifications in 10 mM KCl, similar to wild-type (B)

Overall G4 formation and stability of 5-hmC modifications was examined by ECD as compared to the wild-type, each in the presence of 10 mM KCl alone. $VEGF_{1-4}$ forms a prominent parallel structure as examined by the positive Cotton effect at 265 nm (Figure 3-1B) (Agrawal et al. 2013; Guo et al. 2008; Sun et al. 2005; Sun et al. 2008). Thermodynamic stability is conserved with each modification present individually compared to wild-type (1-4: 67 °C, Mod1: 68 °C, Mod2: 68 °C, Mod3: 67 °C). Further studies were performed to challenge G4 formation when epigenetic modifications are present such as using select co-solvents to mimic physiological conditions outside of cells. Appropriate co-solvent concentrations were determined by depicting the minimal concentration necessary (0-40%) to induce maximal G4 formation through analyzing ECD spectra of $VEGF_{1-4}$ (Figure 3-2). The concentrations of, 10 mM KCl, and 30% MeCN or 10% PEG, were chosen to investigate modified G4 formation. We observed the effect of these co-solvents on G4 stability with epigenetic modifications present compared to wild-type (Figure 3-3A). $VEGF_{1-4}$ thermal stability increased from 67 °C to 83 and 77 °C upon addition of MeCN and PEG, respectively. A knockout oligonucleotide, $VEGF_{KO}$, with G>T mutations in each guanine run within the wild-type sequence was used as a negative control; disruption of G4 formation was confirmed by ECD in the presence of 100 mM KCl with and without 40% MeCN (data not shown). Similar to wild-type, the presence of co-solvents each increased the stability of G4s containing 5-hmC modifications as well. Mod1 was enhanced from 68 °C with 10 mM KCl to 80 °C with MeCN and 77 °C with PEG (Figure 3-3A). $VEGF_{mod2}$ does not form a strong structure, and a parallel orientation is only pronounced when KCl, MeCN, or PEG were added (spectra not shown). Melting temperatures increased from 68 to 84 and 74 °C upon MeCN and PEG additions

compared to KCl control, respectively. Mod3, like mod1, increased stability from 67 °C to 80 °C with MeCN and to 74 °C with PEG. Interestingly, upon the addition of 100 mM KCl, mod2 (T_M of 76 °C) is the only modification that decreased melt temperatures compared to wild-type (T_M of 84 °C) (spectra not shown). Thus, 5-hmC modifications existing on cytosines between guanine runs 1 and 2 (mod1) as well as runs 3 and 4 (mod3) contribute to increased structure stability (Figure 3-1A).

EMSA was used to determine inter- vs. intra-molecular G4 structure formation (Figure 3-3B). The knockout sequence mentioned above was used here as a linear reference in order to indicate G4 migration patterns. Similar to $VEGF_{1-4}$, the 5-hmC-modified G4s displayed downward shifts from the knockout reference, KO, suggesting the existence of intramolecular species. Further migration is seen when 100 mM KCl is added, however, the shift is not marked coinciding with ECD data previously described. When 40% MeCN is added, slower migration of each sequence displays banding above the corresponding control and 100 mM KCl samples. $VEGF_{mod1}$ upon 100 mM KCl and 40% MeCN addition exemplifies intermolecular structures which may be correlated to the decrease in slope of the thermal stability profile shown in Figure 3-3A. This decrease in slope signifies an increase in the number of isoforms existing which may include a cohort of inter- and intra-molecular G4 species combined. Besides mod1, no other sequence allowed for intermolecular species. Instead they closely resembled wild-type. Collectively, it can be concluded that 5-hmC modifications existing on cytosines involved in loop formation do not impair the stability of G4s.

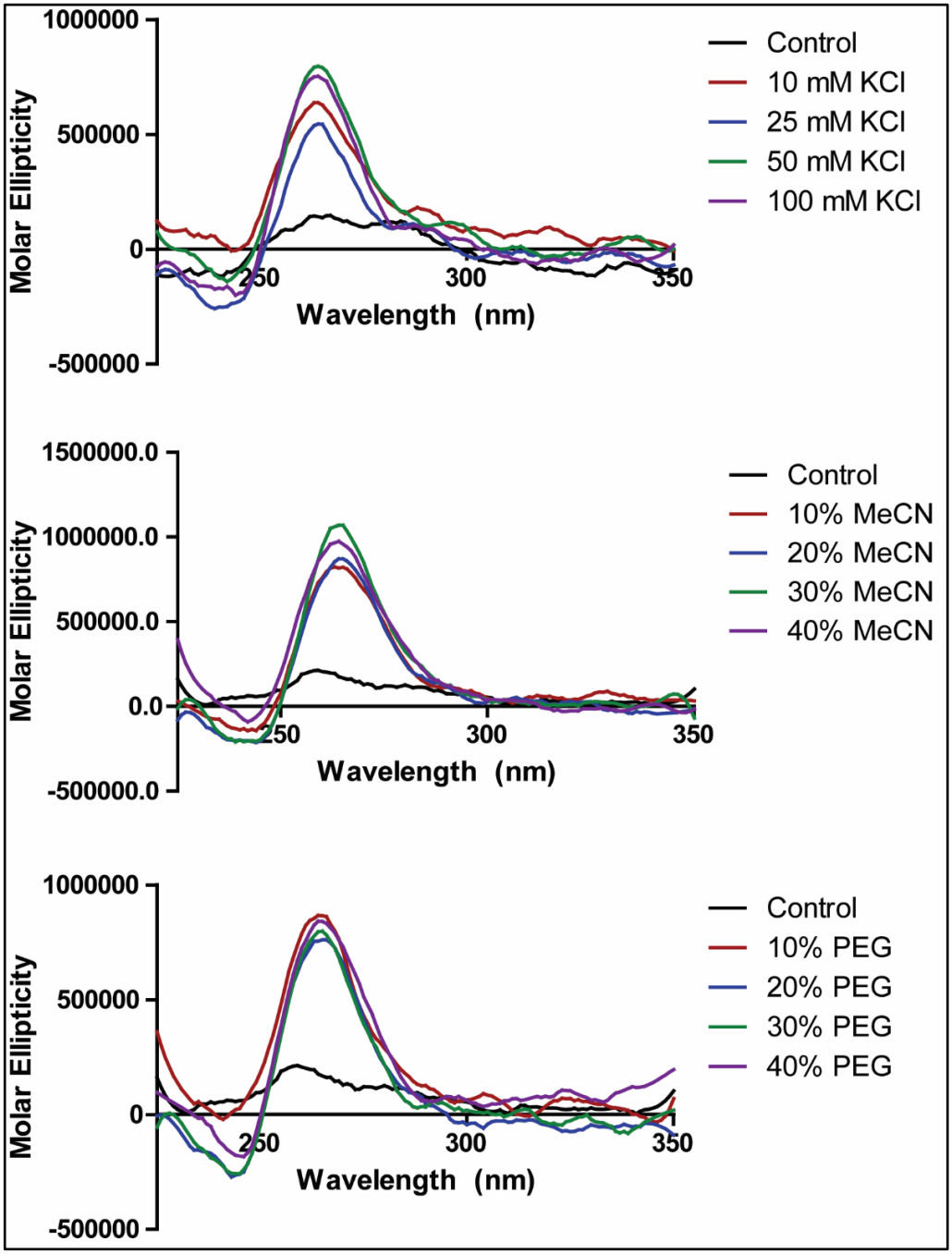


Figure 3-2. Determination of co-solvent concentrations needed for maximum $VEGF_{1-4}$ formation. 0-100 mM gradient of KCl was added to wild-type VEGF G4 to determine 10 mM KCl to be the minimal concentration affecting G4 formation markedly. Similarly, 0-40% of MeCN or PEG gradients, in 10% increments, were added to wild-type resulting in 30% MeCN and 10% PEG being identified for further study.

It was then questioned how these epigenetic modifications impact interactions between G4-stabilizing small molecules and the structure. To date, there are two hypotheses on how small molecules stabilize G4s. First, it is believed that compounds envelope the G4 by binding above and below the structure, involved in π-π stacking, and intermingling with G4 loops. The second hypothesis suggests G4-stabilizing agents intercalate the G4 between tetrads. Because the first has been more widely adopted, we investigated whether 5-hmC modifications may hinder small molecule G4 drug discovery approaches.

Utilizing the universal G4-stabilizing cationic porphyrin, TMPyP4 (Figure 3-4A), for increased G4 thermal stability, any changes were monitored by ECD (Figure 3-4B). 10 μM TMPyP4 (2 equivalents to DNA) was used to allow possible stacking above and below the VEGF G4 with and without modifications present. TMPyP4 increased wild-type and mod1 G4 stability by approximately 12 °C. Compared to wild-type, mod1 had enhanced stability by 7 °C in the presence of TMPyP4. This was not the case for mod2 or mod3 which did not change compared to their respective controls. This suggests that epigenetic modifications can influence G4 drug discovery efforts using small molecules.

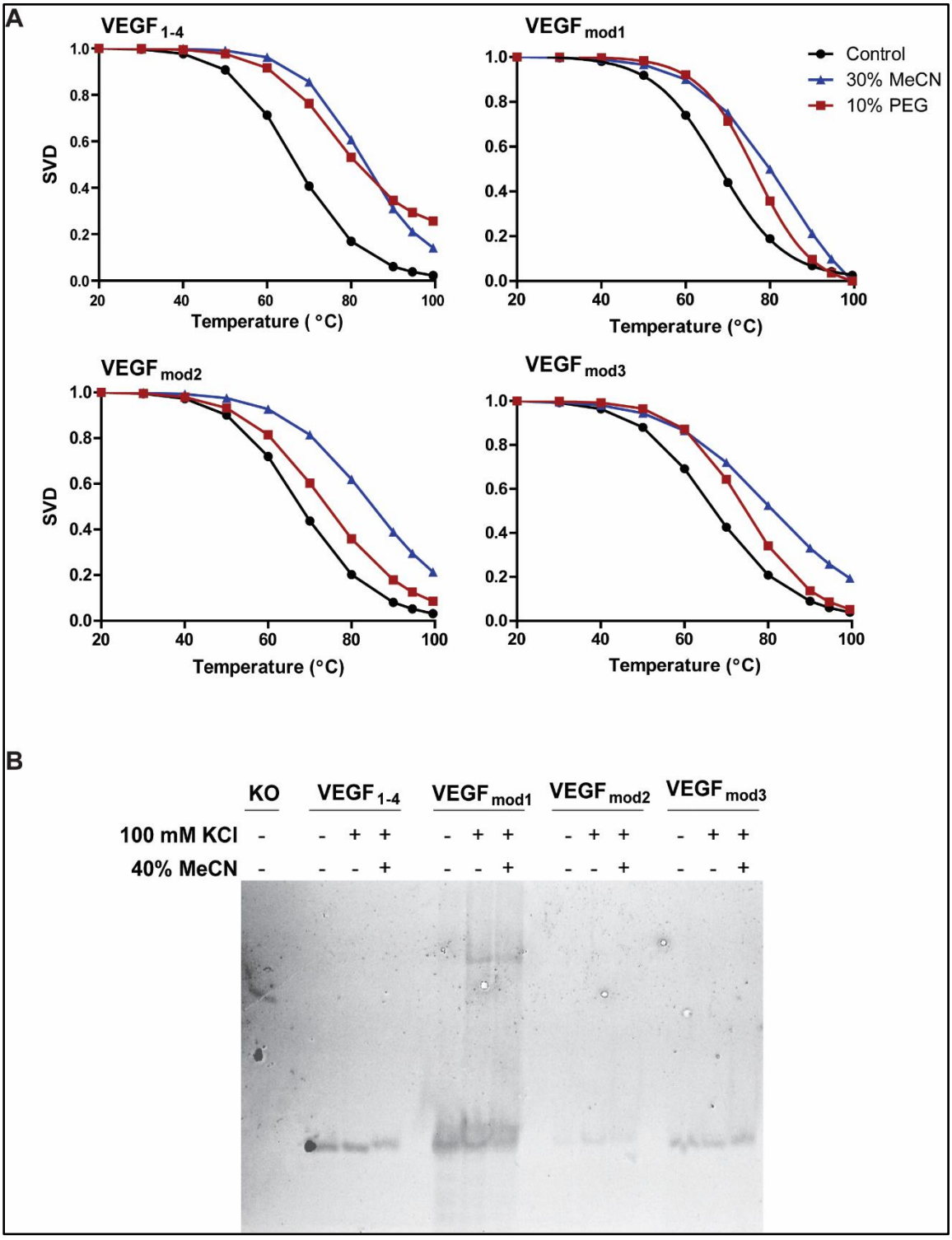


Figure 3-3. Co-solvents mildly increase thermodynamic stability of modified intramolecular VEGF structures. 30% MeCN and 10% PEG increase thermal profiles of each modification as compared to their own control (10 mM KCl) (A) EMSA shows decreased isoforms in wild-type VEGF than with modifications. Mod1, unlike the others, allows intermolecular G4 formation, as well as, intramolecular (B)

Table 3-2. T_M of VEGF modification sequences.

	$VEGF_{1-4}$	**$VEGF_{mod1}$**	**$VEGF_{mod2}$**	**$VEGF_{mod3}$**
10 mM KCl	67	68	68	67
10 mM KCl + 30% MeCN	83	80	84	80
10 mM KCl + 10% PEG	77	77	74	74

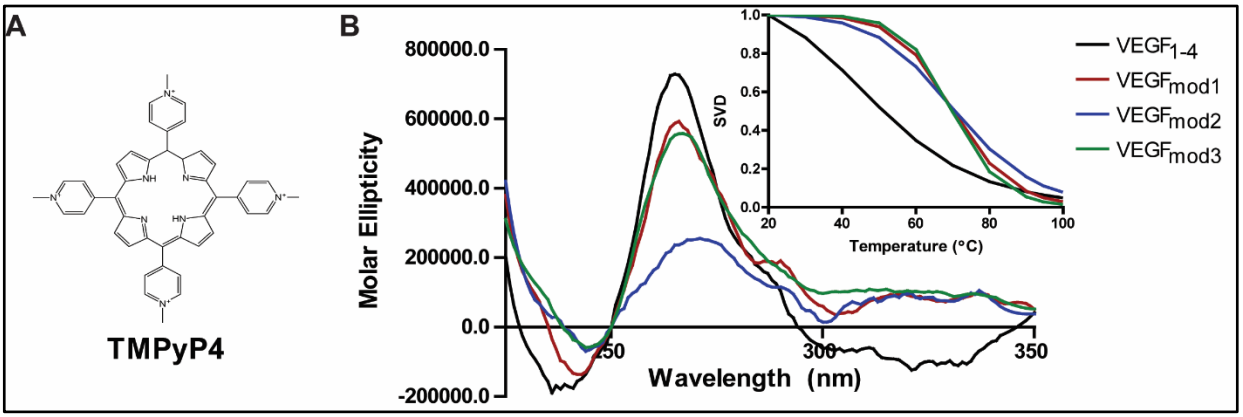


Figure 3-4. 5-hmC modifications increase small molecule effect on G4 thermal stability. Universal G4-stabilizing agent, TMPyP4 (10 μM) (A), enhances VEGF G4 formation, decreases the number of isoforms, and increases thermodynamic stability when epigenetic modifications are present on nucleotides involved in loop formation (B)

Table 3-3. T_M of VEGF modifications with TMPyP4.

	$VEGF_{1-4}$	**$VEGF_{mod1}$**	**$VEGF_{mod2}$**	**$VEGF_{mod3}$**
10 mM KCl	63	70	71	70
10 mM KCl + 10 µM TMPyP4	75	82	71	71

3.3. Using Co-Solvents to Mimic Physiological Conditions for Drug Discovery Screening

Proto-oncogene MYC is amplified in approximately 80% of cancer cases (cBioPortal) (Dang et al. 2006). Its protein functions as a transcription factor and is responsible for regulating more than 15% of genes within the human genome (Chung & Levens 2005). It plays a significant role in cell proliferation, growth, differentiation, and apoptosis (Eisenman 2001). Thus, its deregulation frequently results in tumorigenesis. The c-myc promoter contains a nuclease hypersensitivity element ($NHEIII_1$) upstream from two proximal promoters that initiate 100% of gene transcription (Siddiqui-Jain et al. 2002; Seenisamy et al. 2004). This element contains a guanine dense sequence of 31 bases that can form a G4. This predominantly parallel G4 can act as a silencer of transcription and has been a popular target for G4 drug discovery programs (Figure 3-5) (Brooks & Hurley 2010; Brooks & Hurley 2009). However, with many failed classes of G4-stabilizing compounds, the current G4-stabilizing agent screening approaches need to be optimized by considering the intracellular conditions of a cell's nucleus. Upon consideration, the conditions can be mimicked outside of a cell and applied within screening to best predict the small molecule/G4 DNA interactions occurring in biological systems.

Through the use of co-solvents, physiological mimicry can be accomplished outside of cells. The nature of a nucleus being dehydrated and crowded can be achieved with a myriad of osmolytes and molecular crowding agents. To investigate this use of co-solvents, a small cohort of co-solvents was chosen to add to MYC G4 oligonucleotide sequences (Table 3-2) prior to G4 induction. Osmolytes MeCN, glucose, sucrose, and

ethylene glycol along with their corresponding polymers and molecular crowding agents PEG, dextran sulfate, glycerol, and ficoll were utilized. Preliminary studies with these co-solutes weeded out ethylene glycol, as it prevented stable, inducible G4 formation as observed with ECD (data not shown). However, no other problems arose. Along with co-solvent use, we incorporated another way to mimic nuclear conditions more ideally by extracting nucleoplasm from MiaPaCa-2 cells. This extracted nucleoplasm contains all co-solutes within the nucleus existing outside of DNA and proteins. Each condition was first used to examine any effect on MYC G4 thermal stability by ECD (Figure 3-6A & B).

Three concentrations of extracted nucleoplasm were used to induce MYC G4 formation under relevant physiological conditions (Figure 3-6A). The highest concentration (10%) increased the number of isoforms existing as seen by an increase in slope of its melting profile, but it also shifted the spectra from parallel G4 formation to favoring more ssDNA forms (maxima at 270 nm). This effect was not seen in the presence of 2 or 5% nucleoplasm as these greatly resembled 10 mM KCl control. Similarly, we compared co-solvents to 10 mM KCl control (Figure 3-6B). The concentrations for co-solvents were determined by examining the ECD spectra for the minimum concentration of co-solvent needed to achieve maximal G4 formation as described earlier with VEGF (data not shown). The final concentrations used were 20% MeCN, 20% PEG, 20% ficoll, 10% dextran sulfate, 10% glycerol, 10% glucose, and 20% sucrose. Of these, we observed likenesses between all co-solvents except PEG and MeCN which were destabilizing by 9 °C and stabilizing by 14 °C, respectively. Along with resembling control, these co-solvents were comparable to 2 and 5% nucleoplasm addition as each decreased the number of isoforms existing compared to control as shown by an increase in slope

(ficoll by 74%, dextran sulfate by 78%, sucrose by 48%, glucose by 68%, and glycerol by 48%). Knowing this, the influence on inter- vs intra-molecular isoforms predominating was next evaluated.

A knockout G4 oligonucleotide was used as a linear reference for EMSA (Figure 3-6C). This sequence contains G>T mutations in each of the guanine runs (I-IV) involved in G4 formation. Upon KCl, nucleoplasm, or co-solvent addition, the prime species existing were intramolecular isoforms based on the lower migration patterns. Crisp banding was observed in 2 and 5% nucleoplasm samples compared to 10 mM KCl control, and the co-solvents that best mimicked this were MeCN, dextran sulfate, ficoll, glycerol and glucose. There are greater amounts of isoforms when 10% nucleoplasm, PEG, and sucrose are present. Cooperatively, this suggests that 2 or 5% nucleoplasm, MeCN, dextran sulfate, ficoll, glycerol, and glucose can mimic physiological conditions of a cell's nucleus and limit the amount of G4 isoforms existing. With this, we then wanted to use them in a blind small molecule drug discovery screen.

Förster, or fluorescent, resonance energy transfer (FRET) melt was utilized with a DNA probe containing a fluorophore on the 5' end and a quencher on the 3' end. This DNA probe consisted of the MYC G4 forming sequence, and upon G4 induction, the fluorescence signal is quenched as the fluorophore and quencher are closer together as opposed to its linear form. Compared to a buffer control, stronger G4 thermal stability is indicated by a decrease in fluorescence, an increase in G4 formation, and an increase in melting temperature before returning to its linear state. This is performed blind in 96-well plates and is normally executed under simple buffer or KCl conditions.

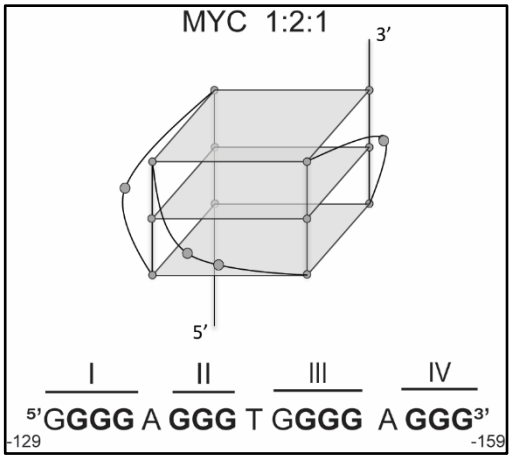


Figure 3-5. The biologically relevant MYC G4. This model depicts a tri-stacked parallel G4 containing guanine runs I-IV and harboring 1:2:1 loop lengths from 5'→3' (Adapted from Brooks & Hurley 2010).

Table 3-4. MYC oligonucleotide sequences.

MYC_{KO}	5’GCGCTTATGTTGAGTGTGTTGAGTGTGTTGAAGGTGTTGAGGAGAC3’
MYC_{KO} EMSA	5’[6~FAM]GCGCTTATGTTGAGTGTGTTGAGTGTGTTGAAGGTGTTGAGGAGAC3’
MYC	5’GCGCTTATGGGGAGGGTGGGGAGGGTGGGGAAGGTGGGGAGGAGAC3’
MYC EMSA	5’[6~FAM]GCGCTTATGGGGAGGGTGGGGAGGGTGGGGAAGGTGGGGAGGAGAC3’
MYC FRET	5’[6~FAM]GCGCTTATGGGGAGGGTGGGGAGGGTGGGGAAGGTGGGGAGGAGAC[TAMRA~6~FAM]3’

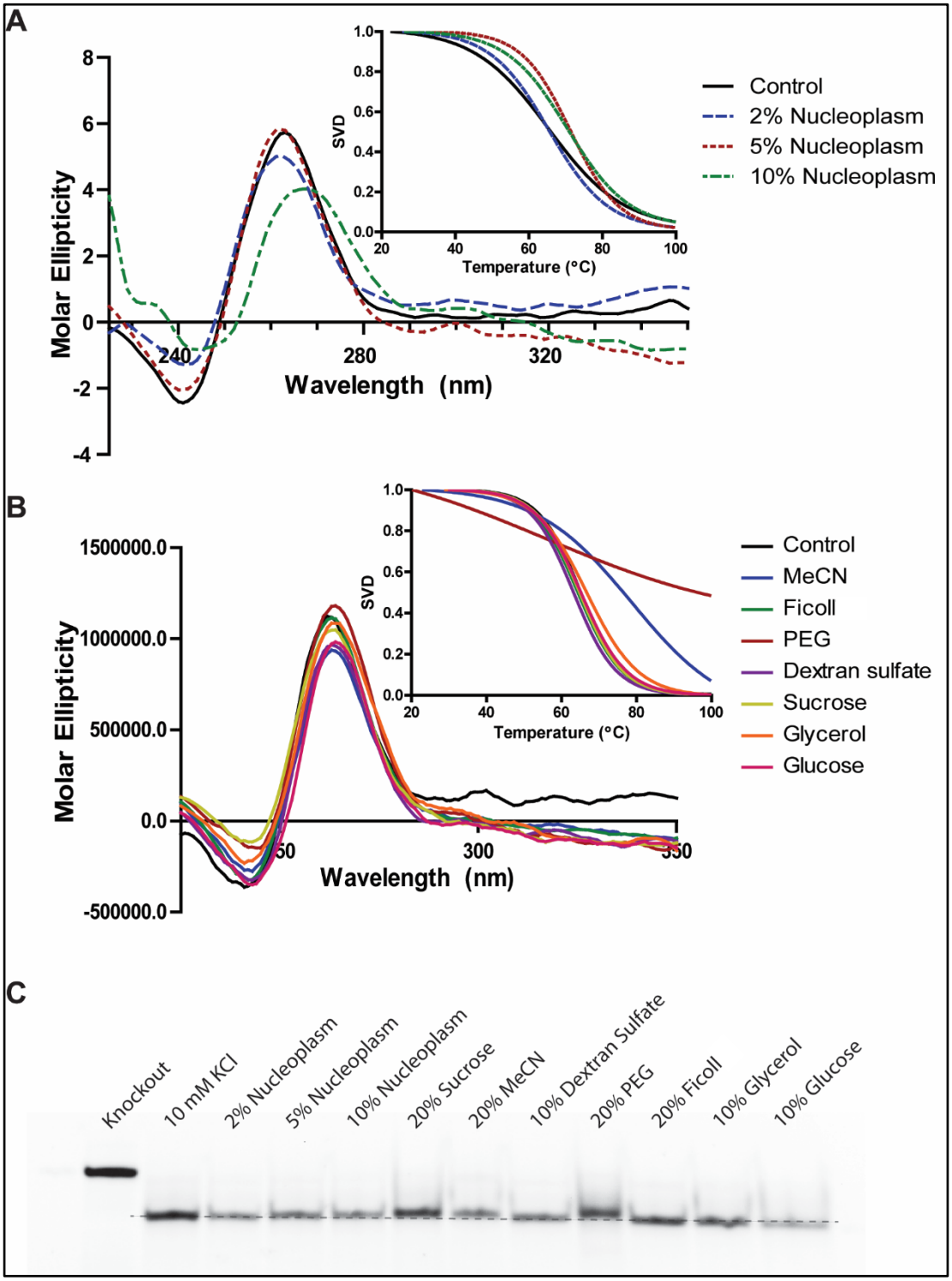
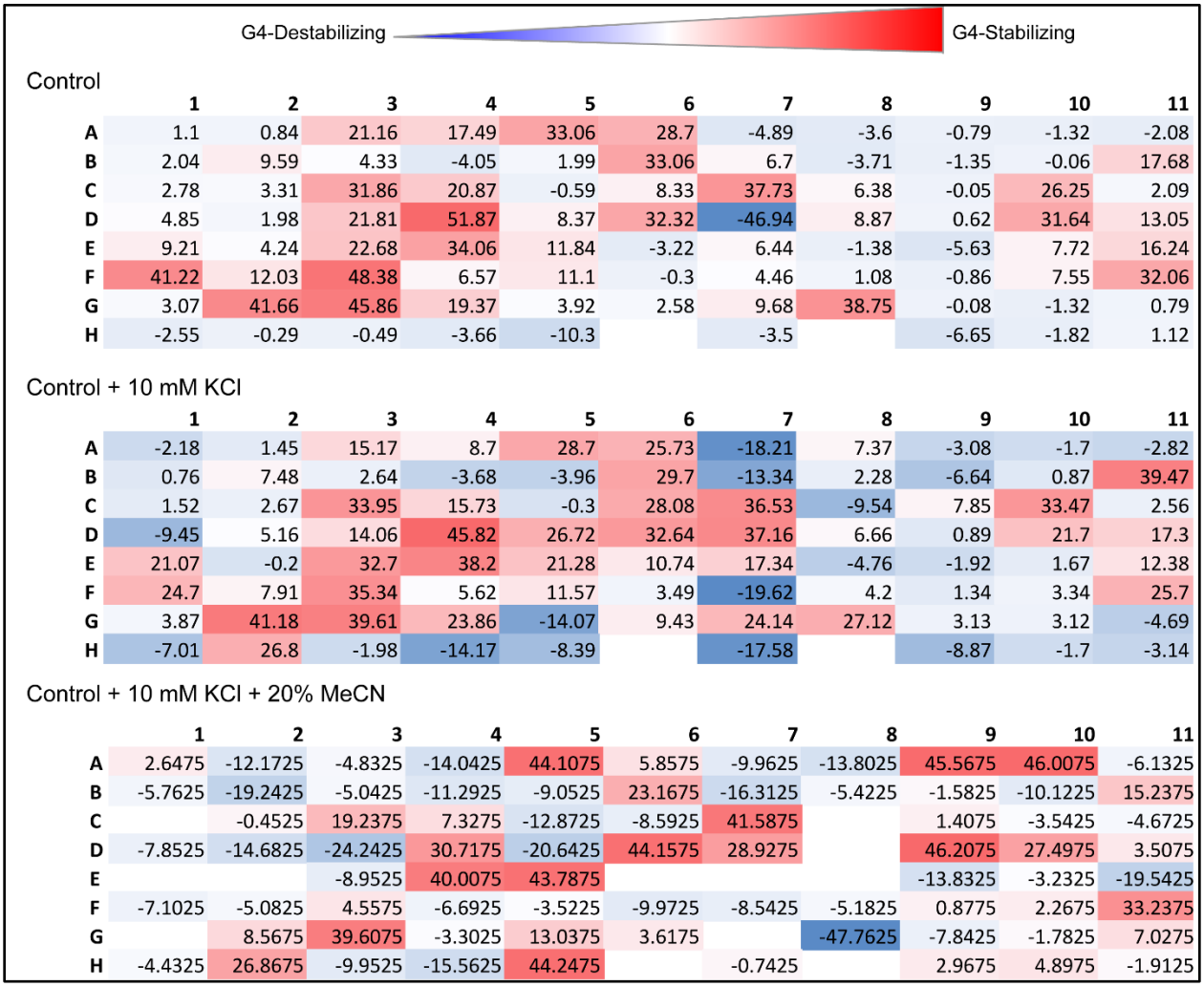


Figure 3-6. Co-solvents and extracted nucleoplasm to induce physiologically relevant average G4 topology. 2 and 5% nucleoplasm stabilizes and decreases the number of G4 isoforms existing (A) most co-solvents are similar to control except PEG and MeCN which destabilize and stabilize the G4, respectively (B) EMSA depicts intramolecular G4 formation with each co-solvent approach with decreased isoforms existing with 2-5% nucleoplasm, MeCN, dextran sulfate, ficoll, and glucose (C)

We applied this same method under three separate conditions (buffer, KCl, and KCl+MeCN) using a single 96-well plate of various compounds ranging from known G4 stabilizers and destabilizers, as well as, novel agents synthesized or isolated at the University of Mississippi (Figure 3-7). The hit rate for lead compounds was 36% in the control (buffer) run and decreased to 26% upon 10 mM KCl and 20% MeCN addition. Many differences in G4 stabilizing compounds were seen between each plate, and not all stabilizing agents within the MeCN run were observed under the more simplistic conditions. These lead compounds were validated by ECD analysis, and only two compounds seemed to favor ssDNA forms as opposed to G4 (SN-38 and EC-2014-13) (Figure 3-8). Thus, utilization of co-solvents decreased the hit rate in compound screening and yielded high success in ECD confirmation. Future studies will investigate cellular effects and the mechanism of action of these lead compounds described.



Control

	1	2	3	4	5	6	7	8	9	10	11
A	1.1	0.84	21.16	17.49	33.06	28.7	-4.89	-3.6	-0.79	-1.32	-2.08
B	2.04	9.59	4.33	-4.05	1.99	33.06	6.7	-3.71	-1.35	-0.06	17.68
C	2.78	3.31	31.86	20.87	-0.59	8.33	37.73	6.38	-0.05	26.25	2.09
D	4.85	1.98	21.81	51.87	8.37	32.32	-46.94	8.87	0.62	31.64	13.05
E	9.21	4.24	22.68	34.06	11.84	-3.22	6.44	-1.38	-5.63	7.72	16.24
F	41.22	12.03	48.38	6.57	11.1	-0.3	4.46	1.08	-0.86	7.55	32.06
G	3.07	41.66	45.86	19.37	3.92	2.58	9.68	38.75	-0.08	-1.32	0.79
H	-2.55	-0.29	-0.49	-3.66	-10.3		-3.5		-6.65	-1.82	1.12

Control + 10 mM KCl

	1	2	3	4	5	6	7	8	9	10	11
A	-2.18	1.45	15.17	8.7	28.7	25.73	-18.21	7.37	-3.08	-1.7	-2.82
B	0.76	7.48	2.64	-3.68	-3.96	29.7	-13.34	2.28	-6.64	0.87	39.47
C	1.52	2.67	33.95	15.73	-0.3	28.08	36.53	-9.54	7.85	33.47	2.56
D	-9.45	5.16	14.06	45.82	26.72	32.64	37.16	6.66	0.89	21.7	17.3
E	21.07	-0.2	32.7	38.2	21.28	10.74	17.34	-4.76	-1.92	1.67	12.38
F	24.7	7.91	35.34	5.62	11.57	3.49	-19.62	4.2	1.34	3.34	25.7
G	3.87	41.18	39.61	23.86	-14.07	9.43	24.14	27.12	3.13	3.12	-4.69
H	-7.01	26.8	-1.98	-14.17	-8.39		-17.58		-8.87	-1.7	-3.14

Control + 10 mM KCl + 20% MeCN

	1	2	3	4	5	6	7	8	9	10	11
A	2.6475	-12.1725	-4.8325	-14.0425	44.1075	5.8575	-9.9625	-13.8025	45.5675	46.0075	-6.1325
B	-5.7625	-19.2425	-5.0425	-11.2925	-9.0525	23.1675	-16.3125	-5.4225	-1.5825	-10.1225	15.2375
C		-0.4525	19.2375	7.3275	-12.8725	-8.5925	41.5875		1.4075	-3.5425	-4.6725
D	-7.8525	-14.6825	-24.2425	30.7175	-20.6425	44.1575	28.9275		46.2075	27.4975	3.5075
E			-8.9525	40.0075	43.7875				-13.8325	-3.2325	-19.5425
F	-7.1025	-5.0825	4.5575	-6.6925	-3.5225	-9.9725	-8.5425	-5.1825	0.8775	2.2675	33.2375
G		8.5675	39.6075	-3.3025	13.0375	3.6175		-47.7625	-7.8425	-1.7825	7.0275
H	-4.4325	26.8675	-9.9525	-15.5625	44.2475		-0.7425		2.9675	4.8975	-1.9125

Figure 3-7. FRET melt drug discovery application of co-solvent MeCN. The same 96-well plate was screened in control, KCl, and KCl+MeCN conditions to examine any differential hits. This “in house” plate consists of known pan-G4-stabilizing agents, as well as various natural products and molecules isolated or synthesized at The University of Mississippi. The screening rate for positive compounds decreased by 10% from control to KCl+MeCN addition. Lead compounds with G4-stabilizing effects are seen in red while blue wells indicate G4-destabilizing agents. White wells depict compounds with no effect on G4 stability.

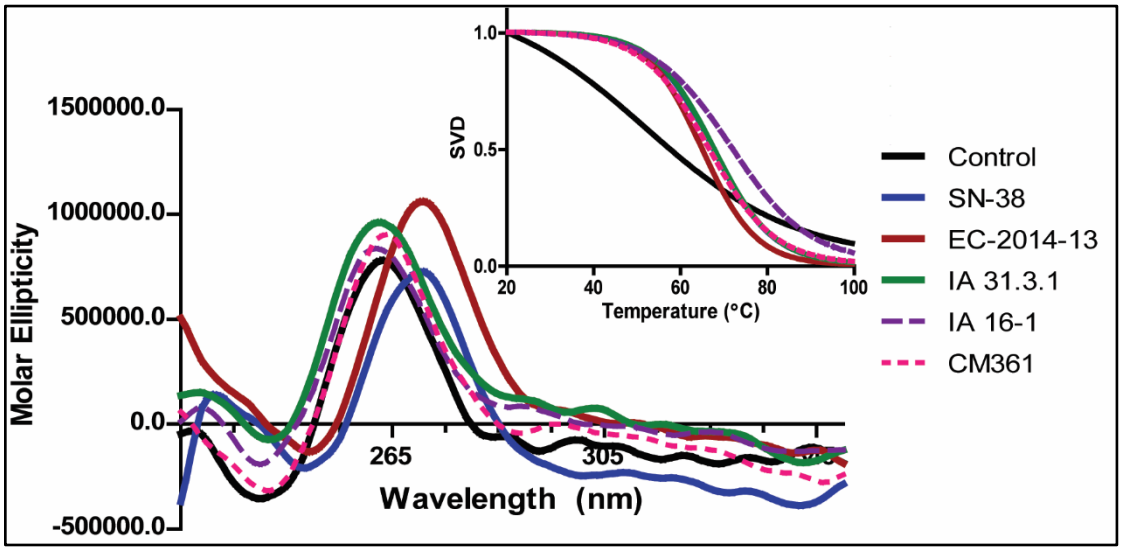


Figure 3-8. Validation of MeCN FRET melt lead compounds. Some compounds (SN-38 and EC-2014-13) shifted G4 formation to ssDNA, but all other compounds increased G4 formation and thermal stability compared to 10 mM KCl control.

3.4. Discussion

G4s have been described to putatively form within promoter and 5' UTR regions throughout the genome, primarily in oncogenes and tumor suppressor genes (Verma et al. 2008; Chambers et al. 2015). This localization around the TSS was conserved amongst four different species (mouse, rat, chimpanzee, and human). In vivo evidence of G4s existing within cells by observing localization of a G4 structure specific antibody, BG4, found G4s within telomeres and non-telomeric regions in nuclei and in cytoplasmic regions—most likely RNA (Biffi et al. 2013). DNA G4s, specifically promoter G4s, have gained much attention for novel targets for anti-cancer therapeutics.

Using G4s as molecular targets for cancer is related to their prevalence within oncogenes and tumor suppressor genes that are highly dysregulated in an array of cancer types. This was found in a genome wide sequence analysis in 2015, but even amongst the initial characterization of promoter G4s within MYC, cancer therapeutics was a central focus. The two most studied promoter G4s are MYC and VEGF, which both form tri-stacked structures with parallel loop configurations. MYC is highly deregulated in about 80% of cancers, and VEGF is frequently overexpressed resulting in angiogenesis, tumor cell proliferation, invasion, and metastasis. Many attempts have been made to suppress MYC and VEGF activity independently, but most have failed. Interestingly, each of the G4s in these promoters function as transcriptional silencing elements. Thus, stabilizing these structures could lead to breakthroughs in cancer research.

Most G4 targeting is centralized around small molecules that selectively bind and stabilize specific structures since globular formations are better targets for selectivity compared to dsDNA. Screening for small molecules that stabilize specific G4s is usually

performed by a FRET melt screen. This procedure is implemented under simple conditions containing buffer and the presence of a monovalent cation like potassium with KCl. This practice has led to the identification of several compound classes that act as G4 stabilizers, but the biological application has not yielded high success. To date, one compound succeeded in vitro and initial in vivo models leading towards the clinic. Quarfloxin failed phase II due to high albumin binding. This is but one example of failed agents for G4 drug discovery, as most are not successful in cells and are not dependent upon a G4 mechanism of action. There are various factors that come into play for drug screening outside of cells such as affinity, selectivity, and specificity. We were primarily concerned with the imitation of the molecular conditions within a cell, more specifically, in this case, within a cell's nucleus in order to recapitulate G4s formed in the nucleus. This led to the current application of co-solvents as mimicry agents of biomolecules within a cell's nuclear environment, as well as the investigation of epigenetic modifications that occur within CpG islands of genes whose promoter regions contain G4-forming regions.

Biomolecules influence cellular activities, as they occupy a substantive amount of space within a cell. These biomolecules allow for the dehydrated and condensed environment within nuclei. MeCN and PEG are common co-solutes reported in G4 literature. Some conclude that PEG positively impacts G4 melting behavior by stabilizing the structure directly as opposed to altering the environment (Corradini et al. 2012; Spink & Chaires 1999). Thus, careful consideration needs to be taken when proposing to use PEGs as molecular crowding agents which led to the alternative agents utilized herein (dextran sulfate, ficoll, and glycerol). Similar criticism has been made regarding the use of MeCN, but it has since been accepted to not alter G4 conformations, but to enhance

G4 stability by manipulating the surroundings (Zhou et al. 2016). Herein we included other osmolytes (sucrose, glucose, and ethylene glycol). Additionally, extracted nucleoplasm was introduced within these studies as an ideal example of mimicking physiological conditions, which is a practice that has not been performed within the G4 field to date.

With this array of co-solutes, we determined the optimal concentrations to be used within MYC (and VEGF) in order to understand how biomolecules affect G4 formation in order to improve future drug discovery efforts. We concluded that low amounts of extracted nucleoplasm, MeCN, dextran sulfate, ficoll, glycerol, and glucose can limit the number of intramolecular G4 isoforms existing when compared to simple buffer +/- KCl controls. Collectively, we distinguished that the effect concentration of co-solutes has on G4 formation differs for each promoter structure, as the amounts used with MYC and VEGF were not identical. In other words, the co-solutes that decrease MYC isoforms did not necessarily decrease VEGF isoforms. This is particularly significant when utilizing co-solvents for mimicking physiological conditions outside of cells during drug discovery screens, as conditions cannot be uniformly applied.

In this work, we utilized MeCN during FRET melt screens for lead compounds stabilizing the MYC G4. Using the same 96-well compound plate, we observed a 10% decrease in lead compounds with 10 mM KCl + 20% MeCN added (23%) compared to the buffer control plate (43%) and KCl alone (44%). Additionally, the lead compounds did not each coincide from plate to plate with varying conditions. Interestingly, a previously described MYC G4-stabilizing compound, quindoline i (sample C10, Figure 3-7), increased G4 thermal stability by 26 °C in the control plate, 33 °C in the 10 mM KCl plate, yet destabilized the G4 by about 4 °C in the presence of 10 mM KCl + 20% MeCN. When

first described, quindoline i bound and stabilized the MYC G4 in vitro and expressed anti-cancer activity via decreased cell proliferation and down-regulation of MYC in hepatocellular carcinoma cells (Ou et al. 2007). However, it was later reported that the c-myc repression was not due to the G4-dependent mechanism (Boddupally et al. 2012; Brown et al. 2011). This finding occurred through the use of RAJI and CA46 Burkitt's lymphoma cells, which contain high levels of MYC expression, and the exon test, which validates whether a compound's mechanism is G4 dependent. Briefly, the two lymphoma cells differ by a translocation between the MYC gene and an immunoglobin gene. RAJI cells contain the translocation which retains both exons 1, 2, and 3, and the promoter containing the MYC G4-forming region. On the other hand, CA46 cells lost the endogenous promoter by reciprocal translocation and only contains exon 2, and 3. Therefore, monitoring the expression of exon 1 and 2 in RAJI cells should match, whereas exon 1 is related to the G4 in CA46 cells and exon 2 is driven by the Ig promoter. It is significant that our FRET melt screen findings on quindoline i match the previous literature both in identifying it as a "hit" in simple buffer conditions and by removing it as a "hit" in the nuclear recapitulating conditions. Future studies will evaluate lead compounds (IA 16-1, IA 31.3.1, and CM361) from the 10 mM KCl + 20% MeCN plate consistent for anti-cancer activity in RAJI and CA46 cells, along with being subjected to the exon test.

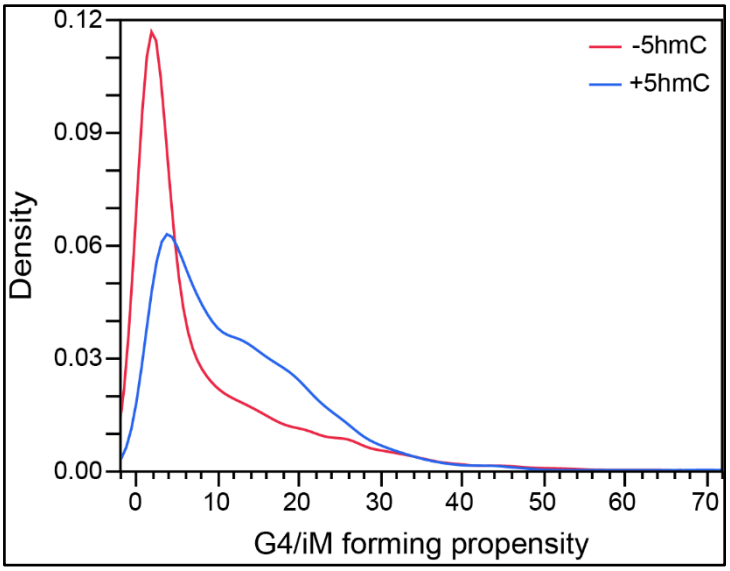


Figure 3.9. G4-forming sequences likely to contain 5-hmC modifications. Fewer 5-mC modifications occur in G4-forming regions while 5-hmC density is co-localized with putative G4s in 100 bp windows centralized around the TSS. There exists a high density of G4s having 5-hmCs downstream from the TSS, in promoter regions. (Adapted from Bhavsar-Jog, Van Dornshuld, Brooks, et al. 2014).

Epigenetic modifications are another set of physiological factors that may affect the stability of G4s. These modifications could be methylation, hydroxymethylation, or acetylation (Bestor 2000). Each have roles in regulating transcription depending on the location and type. Cytosines are the nucleotides susceptible to generating epigenetic modifications, specifically on the 5^{th} carbon atom. Cytosines have a higher chance of being modified when the base is adjacent to a guanine (CpG dinucleotides), commonly termed a CpG island. Such sites are found in almost half of all promoters in the human genome (Taberlay & Jones 2011; Hoffman & Hu 2006; Antequera & Bird 1993). DNA methylation is the most common occurring epigenetic modification and has gained attention for contributing to cancer phenotypes, investigating the roles that methylated CpG islands have during cancer development is critical (Bestor 2000; Taberlay & Jones 2011; Sharma et al. 2009).

Recently, the methylation status on G4s has been analyzed as to how it impacts Bcl-2 promoter activity (Lin et al. 2013). Within the Bcl-2 proximal promoter there lies a CpG island, and methylation of this region led to increased G4 stability which markedly reduced DNA elongation within a PCR stop assay. Lin et. al. claim that methylating the cytosines involved in the Bcl-2 G4 not only increased thermodynamic stability of the structure but also shifted the mixed loop orientation into a more pronounced parallel one. This suggests that the methylated G4 is more likely to keep its conformation rather than unfold which would enhance an inhibitory effect on gene expression. Interestingly, our work suggests that the hydroxymethylation of a CpG island within the VEGF promoter region does not impair or enhance G4 stability but allows the formations of similar stable and inducible parallel G4s as wild-type. However, it is important to note that G4-forming

regions within promoters are more likely to contain 5-hmC than 5-mC modifications (Figure 3-9) (Halder et al. 2010; Bhavsar-Jog, Van Dornshuld, Brooks, et al. 2014).

Previous literature described epigenetic modification-containing CpG islands in induced pluripotent stem cells and endothelial stem cells where they modulate gene expression (Stroud et al. 2011). Endothelial stem cells play key roles in vasculogenesis and angiogenesis, as they form the lining of blood vessels. VEGF is overexpressed by tumors during both processes in order to increase oxygen levels and provide essential nutrients (Turunen & Ylä-Herttuala 2011; Kim et al. 2009; Davis-smyth 2016). It is commonly found upregulated in cancer, mediating tumor growth, invasion and metastasis. Here, we investigated the G4-forming region within the cis-element on the VEGF promoter which has previously been identified as a novel anti-angiogenesis target for cancer therapeutics (Sun et al. 2005; Kim et al. 2016; Agrawal et al. 2013; Sun et al. 2011; Marušič et al. 2013). The G4-forming sequence contains a CpG site, so for drug discovery purposes, it is necessary to explore physiological factors, like epigenetic modifications, that can occur in CpG islands that may affect G4 formation and stability (Yoshida et al. 2016). Thus, G4 stability can be influenced by methylation and possibly hydroxylation depending on the promoter.

With the increasing evidence of G4s existing throughout the genome and within cells, the understanding of the biological function and interactions of secondary DNA structures is necessary to establish (Biffi et al. 2013; Chambers et al. 2015; Huppert & Balasubramanian 2006; Lipps & Rhodes 2009; Murat & Balasubramanian 2014). Such non-canonical structures can affect gene expression, but the regulation behind this function remains in question. Herein, we observed the effects of 5-hmC modifications on

the VEGF G4 and learned that such modifications do not impact overall G4 formation but alters stability in the presence of a stabilizing compound. This suggests that drug discovery efforts should consider the presence of epigenetic modifications on the G4 of interest to consider all potential biological interactions. Also, future studies are necessary to investigate whether protein or transcription factor binding to G4s containing modifications is impeded as is seen with small molecules. Such studies will aid in the understanding of a proposed mechanism behind gene regulation via epigenetic modifications and biomolecules, as well as, further expand G4 drug discovery and research programs to include other disease states like Parkinson's disease, Alzheimer's disease, Fragile X Syndrome, and ALS.

CHAPTER 4. CHARACTERIZING AND TARGETING THE KRAS G4

4.1. Introduction

Over the past decade, pancreatic cancer rates have increased while 5-year survival rates remain steadily in the single digits at around 7% (SEER Cancer Statistics, 2015). This malignant disease accounts for 3% of all cancer incidences within the United States; it can readily grow and metastasize throughout the body, and is likely to avoid detection for a marked amount of time, due to the location of the pancreas (American Cancer Society, 2015). Treatment options are limited as most diagnoses occur well after the cancer has metastasized. A great deal of effort has gone into identifying risk factors associated with the disease, such as smoking, obesity, and diabetes. Unfortunately, none of these considerations have led to earlier detection or have markedly changed survival rates, to date. However, one main factor that researchers are centrally focused on is genetics and the gene mutations, whether inherited or acquired, that may be causative to pancreatic cancer. Out of those being researched, one oncogene seems notable, kRAS.

The RAS gene family consists of five proteins—kRAS, nRAS, hRAS, mRAS, and rRAS—each of which are responsible for encoding intracellular membrane-bound G-proteins involved in cell growth and apoptosis through the Raf-MAP kinase pathway (Lavrado et al. 2015; Saad et al. 2014). The RAS genes are structurally similar and share high homology (approximately 85%) in exon sequences (Friday & Adjei, 2005). Classical RAS proteins—kRAS, nRAS, and hRAS—are located on the short arm of their respective

chromosomes 12, 1, and 11 while rRAS and mRAS are located on the long arms of human chromosomes 19 and 3, respectively. Normally, the encoded $p21^{RAS}$ protein is in one of two states; a GDP-bound inactive state and a GTP-bound active state. In its active state, RAS stimulates the association of RAF to the membrane where kinases can then phosphorylate and activate RAF. Following this event, RAF activates MEK which, in turn, initiates ERK to signal for other substrates allowing for the regulation of cell growth and survival (Wilson & Tolias 2016).

A single point mutation can convert the RAS genes into an oncogenic form, resulting in $p21^{RAS}$ locked into its active form causing aggressive cell growth, metastasis, and anti-apoptosis. Mutations within RAS occur most often at residue 12, 13, or 61, and are prevalent in oncogenesis. RAS protein, rRAS, does not cause malignant transformations like the others in the subfamily (Vähätupa et al. 2016). The most common mutation for kRAS occurs at residue 12 and is seen in 30% of all cancers (majority being solid malignancies) including lung, pancreatic, and colon (Lowy & Willumsen 1993; Ostrem & Shokat 2016). Mutation of nRAS at position 61 is often seen in melanoma and myeloid leukemia, mutants of mRAS typically occur on residues 22 or 71 and are found in breast cancers, and, lastly, frequently expressed in bladder cancer, hRAS mutations happen at either residue 12 or 61 (Ward et al. 2004).

In pancreatic ductal adenocarcinomas alone, mutated kRAS is overexpressed in over 90% of patients (Bailey et al. 2016, cBioPortal for Cancer Genomics). To date, many approaches to suppress mutated RAS protein activity have been tried, but no therapeutic option has yielded specific inhibition in clinical trials. If kRAS expression could be suppressed, then cancerous cell proliferation that is addicted to mutant protein activity

would cease, and the cell could no longer evade apoptosis (Lavrado et al. 2015). This makes the approach of regulating gene expression through transcriptional control an attractive therapeutic focus.

Many attempts have been made to inhibit mutant Ras signaling. However, the approaches tried in these trials failed pre-clinically and clinically. Decreasing levels of GTP-bound RAS seemed promising in order to disrupt the constitutively active RAS protein. However, RAS has high affinity for GTP (approximate $K_d = 10^{-11}$ M) and there is a high concentration of GTP within the cell (approximately 0.5 mM), which makes substrate competition improbable (Ostrem & Shokat 2016). Antisense approaches to block translation of the RAS protein decreased tumorigenicity within mice, but the oligonucleotides lacked a safe and effective drug delivery when applying to human patients (Gray et al. 1993). Some tried to transport the antisense oligonucleotides in viral vectors or plasmids, but there was still no success in tumor specific drug delivery (Zhang et al. 1993; Aoki et al. 1997).

Other studies focused on inhibiting kRAS membrane localization or targeting the protein's downstream effectors. Since Ras proteins are small and hydrophilic, they must undergo four stages of post-translational modifications in order to bind to the intracellular membrane (Ghobrial & Adjei 2002; Adjei 2001; Cho & Lee 2002). Researchers found that farnesyl protein transferase inhibitors block the first step of Ras modifications, thus blocking membrane localization (Queneau et al. 2001; Rowinsky et al. 1999; End 1999). Currently, the approach having the best success involve inhibition of RAS effector signaling. Raf serine/threonine kinases are critical drivers of Ras-mediated oncogenesis through the interactions and activation of MEK1/2 and ERK1/2. Rigosertib is in phase III

clinical trials and works by indirectly inhibiting RAS through blocking the RAS-binding domains of RAS effectors (Ostrem & Shokat 2016; Athuluri-Divakar et al. 2016; Cox et al. 2015). This compound shows great promise in vitro, but there may be issues with efficacy in patients.

Besides working on the Raf-MEK-ERK pathway, others have looked at inhibiting parallel pathways like PI3K-AKT-mTOR. PI3K inhibition has previously shown partial regression of tumor growth in a lung cancer murine model harboring upregulated kRAS (Ebi et al. 2011; Castellano et al. 2013; Gupta et al. 2007). However, other evidence suggests that this pathway is not a significant effector for RAS signaling (Engelman et al. 2008). Though approximately 53 PI3K-AKT-mTOR inhibitors and 29 Raf-MEK-ERK inhibitors are currently in clinical trials, these small molecules are not promising as a monotherapy for cancers harboring mutant RAS (Cox et al. 2015). Preclinical evaluation of possible synergistic activity between these inhibitors may have more promising results (Britten 2013). Many of these anticancer drugs have not passed clinical trials due to their lack of specificity and selectivity for mutant kRAS, as well as, causing the activation of parallel pathways resulting in tumor growth and disease progression (Adjei 2001).

In order to gain specificity and selectivity, immunological approaches have been utilized to differentiate between wild-type and mutant kRAS. Because mutated kRAS can be classified as being tumor specific, a synthetic peptide technique was used to target cytotoxic T-cells to tumors expressing the altered protein where $p21^{RAS}$ already underwent membrane localization (Gjertsen & Gaudernack 1998). Successful peptide delivery resulted in T-cell response to $p21^{RAS}$ and allow the recognition and killing of tumor cells harboring mutated RAS. This posed some difficulties as T-cells require full activation

via multiple signals, and the kRAS mutation was found to be poorly immunogenic which resulted in ineffective vaccinations (Fossum et al. 1995; Gjertsen et al. 2001).

Recently, small molecule approaches to target mutant kRAS via the kRAS-SOS binding pocket has yielded promising results (Wilson & Tolias 2016). Stephen Fesik's research group at Vanderbilt University has synthesized compounds that bind to the GDP-bound kRAS mutant G12D through the use of fragment-based screening (Sun et al. 2012). Lead compounds were shown to weakly bind the kRAS binding pocket initiating a change in protein structure and opening up a new binding pocket. Blocking this binding pocket prevents kRAS-SOS interactions and does not allow for kRAS to initiate its downstream signaling cascade. However, further optimization of these compounds is needed to gain the necessary affinity required to move the compounds into clinical trials.

Another team is looking at small molecules that irreversibly bind to kRAS G12C through direct targeting of the cysteine mutation (Ostrem et al. 2013). The fragments utilized covalently modify the cysteine and bind to the switch-II region on kRAS causing decreased cell viability and increased cell death for lung cancer cell lines harboring the mutation when compared to cells with wild-type kRAS. Unfortunately, the compounds have limited potency to move from preclinical to clinical trials. Based on these approaches, suppressing kRAS expression has shown preclinical promise, but no viable molecular target has yet been established.

To tackle this "holy grail" of cancer therapeutics, we focused on an alternative method to target kRAS involving regulating gene expression through transcriptional control via secondary DNA structures, G-quadruplexes (G4s). This approach will ultimately decrease protein expression and thus halt kRAS protein activity and likely

cause apoptosis of the addicted cancerous cell. This not only is a unique approach to cancer therapeutics, but it will expand promoter G4 drug discovery programs into examining kRAS as a target for pancreatic cancer therapeutics. We hypothesize that unique G4s in the kRAS promoter will have biological functions and may serve as molecular targets for anti-cancer development. To investigate this hypothesis and overall tactic, we focused on two specific aims allowing us to determine the biologically relevant kRAS G4 structure, characterize compounds with the G4-stabilizing potential, identify the major structure present with compound-induced stabilization, and to examine the mechanism of action of such agents in pancreatic cancer cells.

Specific Aim 1. *Elucidate the predominant and biologically relevant G4 isoform within the kRAS promoter region.*

Specific Aim 2. *Identify and characterize kRAS-mid-G4-stabilizing compounds in vitro and in cell culture.*

4.2. Form and Function of Multiple G4s in the kRAS Promoter

The region of the proximal promoter in the kRAS gene from -324 to -39 bp, in relation to the transcriptional start site (TSS), contains a high density of guanine residues (Figure 4-1A). The three distinct guanine runs have potential to form multiple G4 isoforms that can act as silencing elements (Lavrado et al. 2015). One has been previously studied resulting in the finding of numerous isoforms (Cogoi & Xodo 2006; Xodo et al. 2008; Cogoi et al. 2008; Cogoi et al. 2009; Paramasivam et al. 2011). The molecular models of G4 32R portray di- and tri-stacked tetrads with conflicting loop sizes and directionality. The first 2 isoforms characterized were parallel tri-stacked G4s harboring a kinked cytosine loop connecting the first and second guanine runs (Cogoi & Xodo 2006). Under further examination, this model was proposed to contain a thymine in the second run instead as suggested by Cogoi et al. 2008. Next, a model arose containing the kinked thymine in the second guanine run with an inclusion of T/G (Paramasivam et al. 2011; Paramasivam et al. 2009). Further examination of 500 bp upstream of the TSS exposed the previously described G4 and two uncharacterized G4-forming regions.

We refer to the known 32 bp sequence as the near region, capable of forming the $G4_{near}$, sitting -128 bp from the TSS. Labeled based on their proximity to the TSS, the near (-129 to -161), mid (-174 to -226), and far (-238 to -273) regions have been identified as capable of forming G4s based on the consensus pattern $(G_3N_{1-9})_3G_3$ (Figure 4-1A) (Morgan et al. 2016). This work focuses on elucidating the predominant G4s that form within the kRAS promoter, understand any biological role in gene regulation, and targeting such structures with small molecules for pharmacological intervention in pancreatic cancer therapeutics. All oligonucleotides used are described in Table 4-1.

G4 formation and overall thermal stability was assessed by ECD in the absence and presence of G4-stabilizing KCl (100 mM) alone or combined with 40% acetonitrile (MeCN) (Figure 4-1B). $G4_{near}$ formed all parallel configured isoforms, as depicted by the positive Cotton effect at 262 nm in KCl with or without MeCN, which has been formerly observed in literature (Bari & Pescitelli 2010). Thermal stability was increased from 46 °C, in buffer only, to 59 °C with KCl present, and to 69 °C with the addition of MeCN. $G4_{mid}$ forms a mixed parallel and antiparallel structure as examined by two positive maxima at 263 and 290 nm, respectively. Upon KCl addition to buffer only, thermal stability improved from 56 to 82 °C, and increased to >95 °C with the addition of 40% MeCN. Unlike the others, the far region does not form higher order structures, as ECD spectra only notes spectral maxima indication possible parallel and antiparallel conformation with KCl and MeCN combined. Additionally, these G4s did not form strong enough to determine thermal stability.

Total knockout sequences were created for each G4-forming region by G>T mutations (Table 4-1). G4 formation was disrupted as observed by ECD in the presence of 100 mM KCl as predicted (Figure 4-1C). Such sequences were used as linear species controls in EMSA.

Table 4-1. kRAS oligonucleotide sequences.

$G4_{near\text{-}KO}$	5'AGGGCTTTTTGGGAAGAGTGAAGAGGGGGAGG3'
$G4_{near\text{-}KO}$ EMSA	5'[6~FAM]AGGGCTTTTTGGGAAGAGTGAAGAGGGGGAGG3'
$G4_{near}$	5'AGGGCGGTGTGGGAAGAGGGAAGAGGGGGAGG3'
$G4_{near}$ EMSA	5'[6~FAM]AGGGCGGTGTGGGAAGAGGGAAGAGGGGGAGG3
$G4_{mid\text{-}KO}$	5'CGGGGAGAAGGAGTTTGCCGGGCCGGGCCGGCGGGGGAGG AGCGTTTGCCGGGC3'
$G4_{mid\text{-}KO}$ EMSA	5'[6~FAM]CGGGGAGAAGGAGTTTGCCGGGCCGGGCCGGCGGG GGAGGAGCGTTTGCCGGGC3'
$G4_{mid}$	5'CGGGGAGAAGGAGGGGGCCGGGCCGGGCCGGCGGGGGAG GAGCGGGGGCCGGGC3'
$G4_{mid}$ EMSA	5'[6~FAM]CGGGGAGAAGGAGGGGGCCGGGCCGGGCCGGCGG GGGAGGAGCGGGGGCCGGGC3'
$G4_{mid}$ DMS	5'[6~FAM]TTTTTTTCGGGGAGAAGGAGGGGGCCGGGCCGGGC CGGCGGGGGAGGAGCGGGGGCCGGGCTTTTTTT3'
$G4_{mid}$ FRET	5'[6~FAM]GCGGGGAGAAGGAGGGGGCCGGGCCGGGCCGGCG GGGGAGGAGCGGGGGCCGGGC[TAMRA~6~FAM]3'
$G4_{mid}$ DMS Forward Primer	5'-[6~FAM]GATGCGTTCCGCGCTCGA-3'
$G4_{mid}$ DMS Reverse Primer	5'-[6~FAM]AGTCCCTCCTCCCGCCAA-3'
Mut A	5'CGTTGAGAAGGAGGGGGCCGGGCCGGGCCGGCGGGGGAG GAGCGGGGGCCGGGC3'
Mut B	5'CGGGGAGAAGGAGGTGGCCGGGCCGGGCCGGCGGGGGAG GAGCGGGGGCCGGGC3'
Mut C	5'CGGGGAGAAGGAGGGGGCCGTGCCGGGCCGGCGGGGGAG GAGCGGGGGCCGGGC3'
Mut D	5'CGGGGAGAAGGAGGGGGCCGGGCCGTGCCGGCGGGGGAG GAGCGGGGGCCGGGC3'
Mut E	5'CGGGGAGAAGGAGGGGGCCGGGCCGGGCCGGCGGTGGAG GAGCGGGGGCCGGGC3'
Mut F	5'CGGGGAGAAGGAGGGGGCCGGGCCGGGCCGGCGGGGGAG GAGCGGTGGCCGGGC3'
Mut G	5'CGGGGAGAAGGAGGGGGCCGGGCCGGGCCGGCGGGGGAG GAGCGGGGGCCGTGC3'
$G4_{far\text{-}KO}$	5'AAGGTGTGGCTGTTGCGGTCTAGTGTGGCGAGCCGTGCC3'
$G4_{far\text{-}KO}$ EMSA	5'[6~FAM]AAGGTGTGGCTGTTGCGGTCTAGTGTGGCGAGCCGT GCC3'
$G4_{far}$	5'AGGGGTGGCTGGGGCGGTCTAGGGTGGCGAGCCGGGC3'
$G4_{far}$ EMSA	5'[6~FAM]AGGGGTGGCTGGGGCGGTCTAGGGTGGCGAGCCGG GC3'

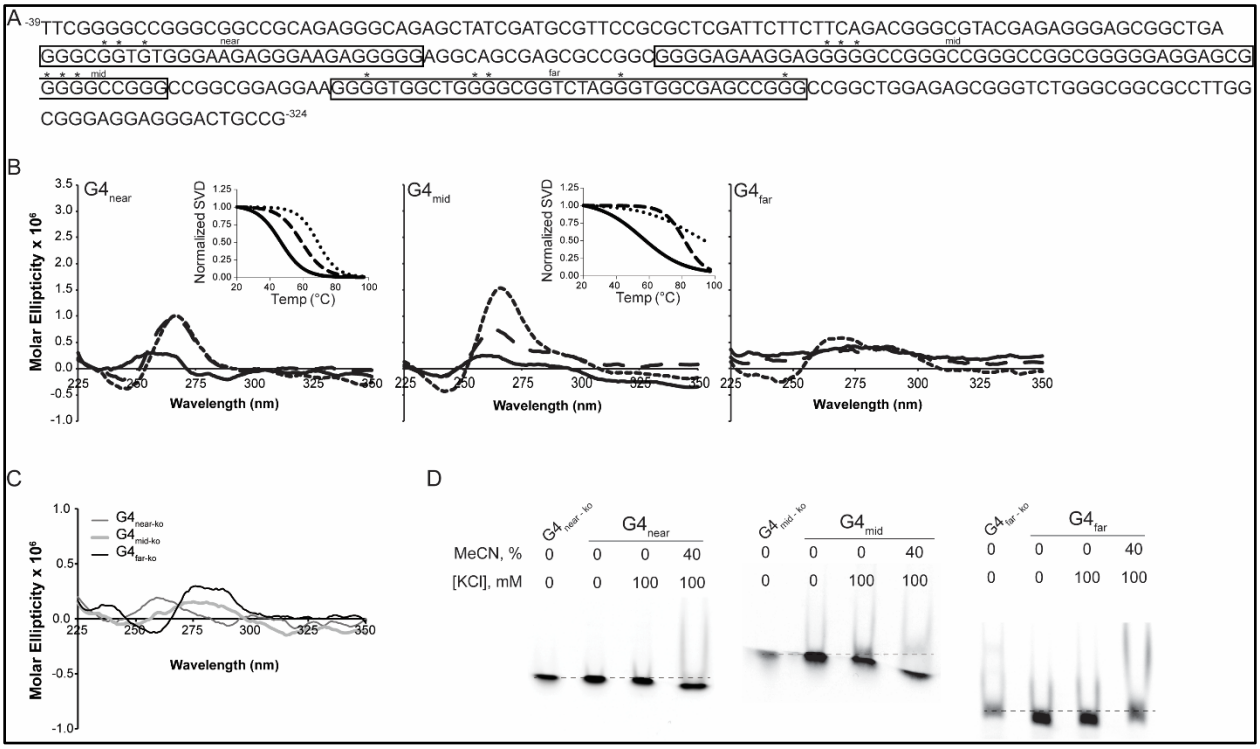


Figure 4-1. Three potential G4-forming regions within the kRAS promoter. The proximal promoter region on the kRAS gene highlighting three G4-forming regions: near, mid, and far (A) ECD spectra of kRAS G4-forming regions in buffer only (solid line) with the addition of 100 mM KCl (long dash line) and with 100 mM KCl + 40% MeCN (short dash line) (B) Total G-to-T mutations (denoted in A by asterisks) prevent G4 formation (C); these mutated sequences were used in EMSA as linear references to show the formation of inter- or intramolecular structures in the presence of buffer alone, 100 mM KCl, or 100 mM KCl + 40% MeCN. Reduced isoforms of intramolecular G4s forming in near and mid regions were observed, especially upon 40% MeCN addition (D) (Adapted from Morgan et.al., 2016)

In order to differentiate between inter- and intra-molecular G4 formations, electromobility shift assay (EMSA) was employed (Figure 4-1D). The total knockout sequences described above were used as reference for the migration of linear species. Faster migration compared to linear strands verifies the formation of compact G4 structures. With KCl present, a downward shift migration is observed for $G4_{near}$ and $G4_{mid}$. An even greater shift occurred with the further addition of 40% MeCN. In all, ECD and EMSA data indicate that the $G4_{far}$ region is not likely to form stable intramolecular G4s, as compared to the stable and inducible structures note in $G4_{near}$ and $G4_{mid}$.

Structure elucidation and confirmation of $G4_{near}$ was performed in Dr. Brooks' laboratory by Dr. Vanessa Gaerig, in order to clarify the predominant G4 (data not shown); several varying structures have been reported previously (Cogoi et al. 2008; Cogoi & Xodo 2006; Paramasivam et al. 2011; Paramasivam et al. 2009; Cogoi et al. 2009). Radioactive DMS footprinting with the 32-nucleotide near kRAS promoter region and two G>T mutant sequences was performed. In the presence of 100 mM KCl, the predominant G4 forming within the wildtype sequence contains three guanine runs that are continuous (A, C, and E) and one discontinuous, or kinked, run (B). G-to-T mutation of run B disrupted G4 formation while mutation of run D preserved the guanine protection pattern noted in wildtype. Collectively, ECD spectra and DMS footprinting supports the formation of a tri-stacked structure harboring a $1{:}1_k{:}1{:}12$ parallel loop conformation (k refers to a single "kinked" thymine in run B between guanines). This data coincides with one of the previously proposed models depicting a kinked higher order structure (Paramasivam et al. 2009).

Unlike $G4_{near}$, $G4_{mid}$ has not been previously described. This 54-nucleotide region within the kRAS promoter consists of seven runs of contiguous guanines, referred to as runs A-G from 5'→3' (Figure 4-2A). DMS footprinting was executed in the absence and presence of 100 mM KCl on the whole wild-type sequence. Banding patterns were described by histogram using ImageJ software, and this evidence suggests several species of $G4_{mid}$ exist in equilibrium with partial protection patterns observed in runs B, E, and F (Figure 4-2B).

Mutant sequences were created within each of the seven guanine runs to help characterize $G4_{mid}$. Each mutated run contained critical G>T mutations in order to disrupt G4 formation when specific runs of guanines are required. We observed such disruption by ECD spectra (Figure 4-3) and thermal analysis in 10 mM KCl, 10mM KCl + 20% MeCN, and 100 mM KCl (Figure 4-2C, Table 4-2). All G4 structures formed maintained the mixed loop orientation (Figure 4-3) similar to that observed in wildtype (Figure 4-1B). Amongst all conditions, G4 destabilization was evident when runs B, E, and F were mutated as thermal stability was markedly decreased in all conditions. Thus, from ECD analysis and DMS footprinting, the major isoform model proposed herein is a tri-stacked structure consisting of runs B, C, E, and F having intervening loops of 2, 10, and 8 nucleotides from 5'→3', respectively (Figure 4-2D).

Table 4-2. T_M of $G4_{mid}$ sequences (Morgan et al. 2016).

	Wild-type	**Mut A**	**Mut B**	**Mut C**	**Mut D**	**Mut E**	**Mut G**	**Mut F**
10 mM KCl	70	65	50	63	65	63	53	65
10 mM KCl + 20% MeCN	73	70	64	71	71	64	64	71
100 mM KCl	81	78	70	89	82	84	63	87

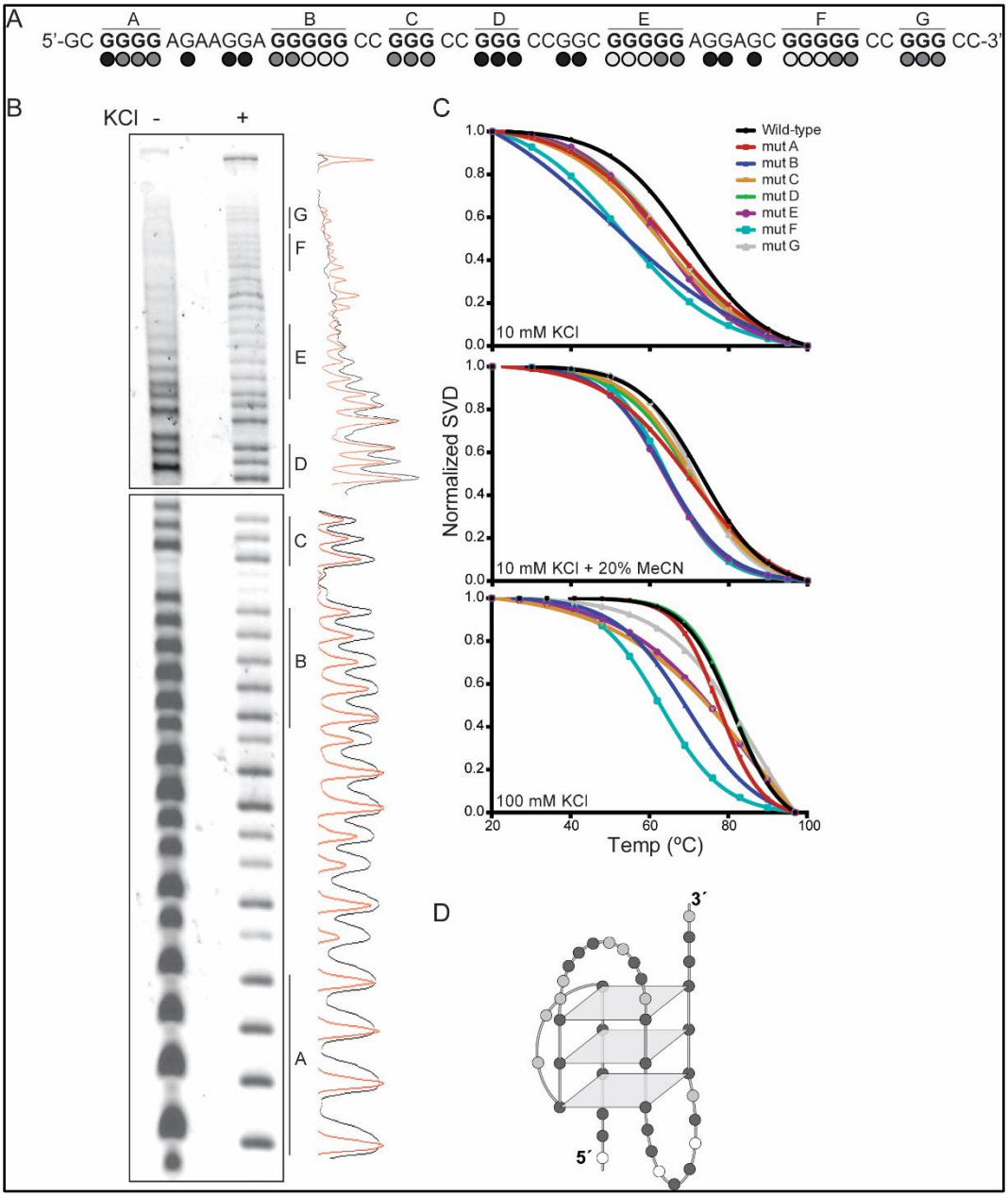


Figure 4-2. Elucidation of $G4_{mid}$ G4 topology. DMS footprinting of $G4_{mid}$ with 100 mM KCl. Full protection is seen in guanine runs B, E, and F; partial protection in runs A, B, C, E, F, and G; no protection seen in run D (A) Full sequence fluorescent DMS sequencing gel with no KCl control and 100 mM KCl (B) thermodynamic stability of various G>T mutations that disrupt guanine runs A-G (C) proposed topology of $G4_{mid}$ as a tri-stacked G4 with runs B, C, E, and F with a 2:10:8 loop configuration in the 5'→3' direction (D) (Morgan et al. 2016).

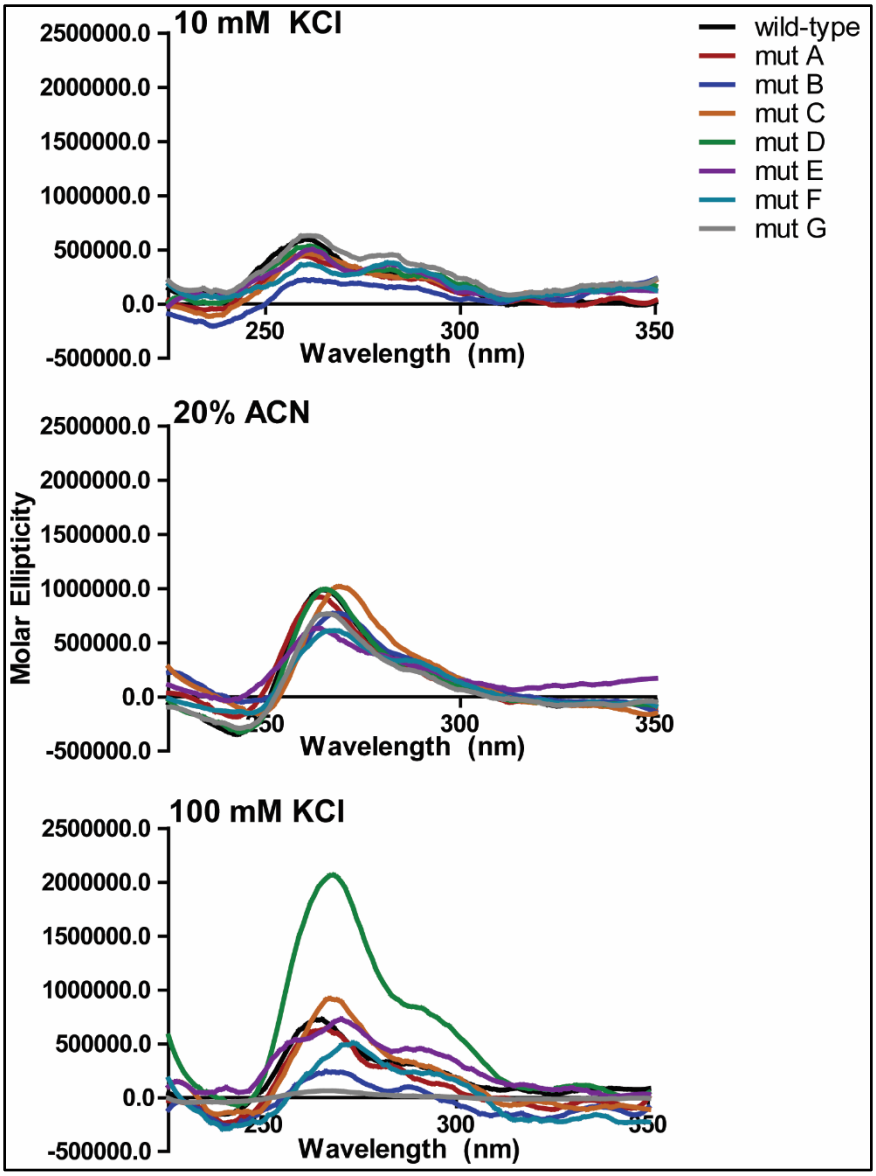


Figure 4-3. G4 topology of kRAS mutated sequences under variant co-solvent conditions. Full G>T mutations were made in each of the seven runs of guanines within the kRAS mid-region sequence on the promoter. ECD spectra showed no effect on topology in any of the mutated sequences compared to wildtype. Each sequence showed a mixed parallel/antiparallel configuration as denoted by positive Cotton effects around 265 and 290 nm, respectively.

Next, we wanted to evaluate the biological function of G4s within the kRAS promoter. Previous literature noted transcriptional silencing of kRAS in the presence of cationic porphyrin and known G4 stabilizing agent, TMPyP4 in Panc-1 (pancreatic adenocarcinoma), HCT116 (colon cancer), and SW620 (colorectal adenocarcinoma) cells (Cogoi et al. 2008; Lavrado et al. 2015). FRET melt studies tested the effect of this compound on thermal stability of both $G4_{near}$ and $G4_{mid}$ (Figure 4-4A). $G4_{near}$ increased from 34 to 65 °C, and $G4_{mid}$ alone was 52 °C heightening to 73 °C. In contrast, we used TMPyP2, a constitutional isomer of TMPyP4, as a negative control because it does not interact with G4s. As predicted, TMPyP2 had no effect on thermal stability for $G4_{near}$ and decreased that of $G4_{mid}$ by 6 °C. To test these compounds' effect on cell survival and, more specifically, the effect on regulating kRAS, we used the pancreatic cell line Panc-1. These cells are mildly addicted to their kRAS mutation and are commonly used in studies on kRAS (Lieber et al. 1975; Cogoi & Xodo 2006; Cogoi et al. 2013; Lavrado et al. 2015). The MTS assay was utilized to measure cell viability in the presence of TMPyP2 or TMPyP4 from 0-500 μM for an incubation time of 48 hr. These compounds have an absorbance at 490 nm which is the same as the formazin measured through the MTS assay. Thus, in order to deduct this colorimetric contribution from the background, we measured absorbance of matched doses of compound before (pre-wash) and after removing (post-wash) and replacing all media from the plates (Figure 4-4B, performed by Dr. Tracy Brooks). Post-wash conditions showed little-to-no effect on cell viability while pre-wash conditions showed a dose-dependent decrease in survival likely due to the colorimetric contribution of the compound and over-estimation of decreased cell viability

post-subtraction (only by approx. 25% with TMPyP4 and no effect with TMPyP2). From this data, we moved forward with compounds at concentrations of 25 and 50 μM.

kRAS mRNA expression was monitored by qRT-PCR in Panc-1 cells upon 25 μM TMPyP4 or TMPyP2 incubation for 48 hr (Figure 4-4C, performed by Dr. Tracy Brooks). Our findings show that TMPyP4 significantly decreased kRAS expression ($p<0.05$) by 45% compared to expression from untreated control. As expected, TMPyP2 did not change kRAS mRNA expression. Thus, this evidence supports the idea that silencing of kRAS expression is mediated by G4 stabilization.

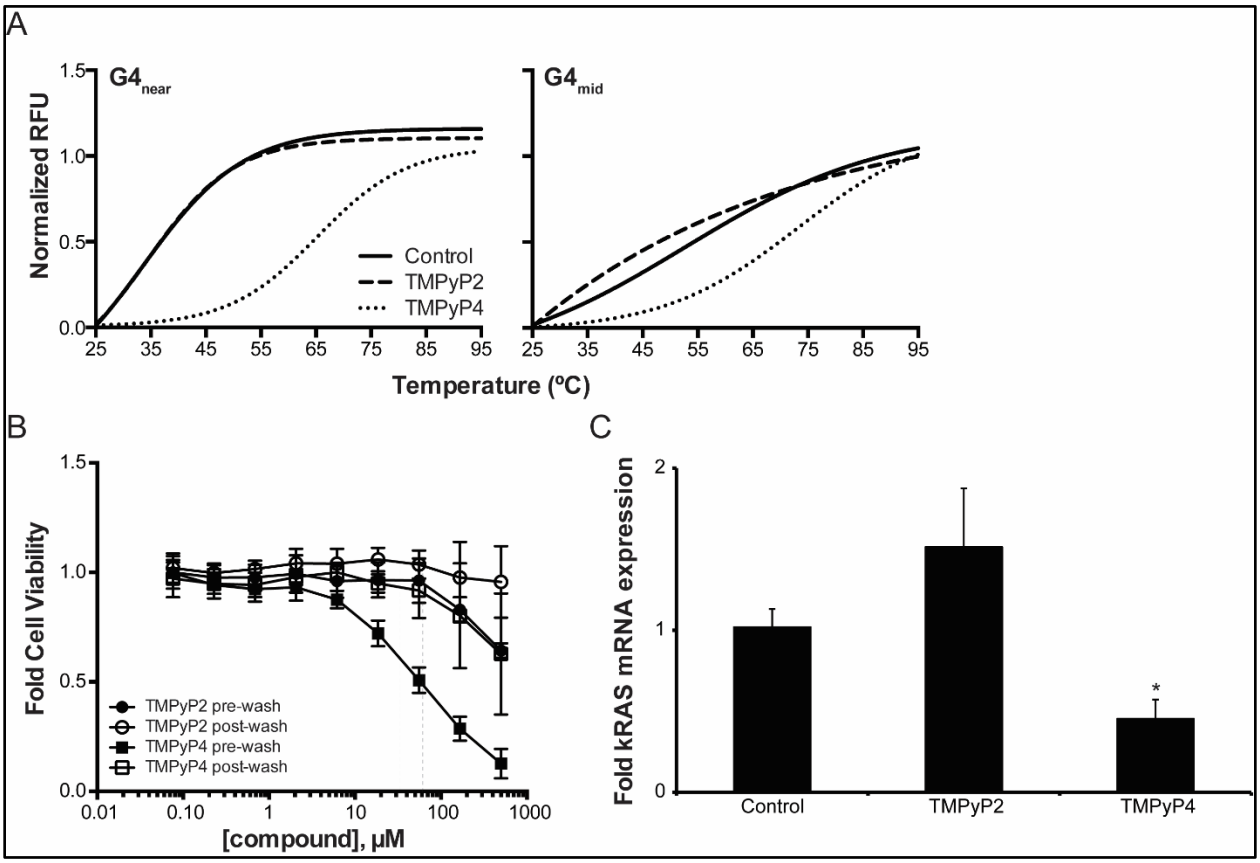


Figure 4-4. TMPyP4-induced G4 stabilization on pancreatic cancer cell viability and kRAS expression. TMPyP4, but not TMPyP2, markedly stabilized $G4_{near}$ and $G4_{mid}$ as examined by FRET melt (A) Cell viability in Panc-1 cells showed decreased survival when compound was not washed (pre-wash) from plate and replaced with fresh media. This can be related to the colorimetric contribution of TMPyP4 causing an over-estimation of reduced cell viability after washing away the compound (post-wash) (B) Panc-1 cells were treated with either 25 µM TMPyP2 or TMPyP4 for 48 hr, and kRAS mRNA expression was examined, as normalization to GAPDH, TMPyP4 significantly (*$p<0.05$, as determined by one-way ANOVA with Tukey post-hoc analysis) decreased promoter activity and, ultimately, kRAS mRNA expression (C) (Morgan et al. 2016).

Next, we wanted to assess which G4 is responsible for the silencing effect on kRAS transcription. A panel of luciferase promoters were constructed to examine the kRAS full promoter and isolated G4 regions (Figure 4-5A). All plasmids were assembled by inserting the regions of interest within the kRAS promoter into the multiple cloning region (MCR) of a promoterless pGL4.17 firefly luciferase vector (EV). The kRAS-500 plasmid was built containing the promoter from -500 to +0 from the TSS amplified by PCR from human genomic DNA which was previously described to inhabit G4 elements (Song et al. 2009; Lavrado et al. 2015). Another plasmid was constructed by Eurofins Operon containing the -324 to +50 bp region, with respect to the TSS. This region has been reported significant for full promoter activity and contains the near, mid, and far G4-forming regions (Jordano & Perucho 1988; Matsuo & Yamazaki 1989; Yamamoto & Perucho 1988). This plasmid was used to derive two mutated plasmids that contained G-to-T mutations in either the $G4_{near}$ (termed 324 mt Near) or the $G4_{mid}$ (termed 324 mt Mid) regions through site-directed mutagenesis; G-to-T mutations were made to each region like those in the knockout oligonucleotides in Table 4-1 and Figure 4-1C. The promoterless, empty vector (EV) and promoter-driven SV40 plasmids were used as negative and positive controls, respectively.

Basal expression was first examined from each plasmid after 48 hr transfection in HEK-293 cells, where firefly activity was normalized to Renilla luciferase expression driven by pRL. The fold RLU for EV activity was 0.03 ± 0.01, SV40 activity was about 50-fold higher than EV with 1.46 ± 0.08, and the kRAS plasmids were somewhere in between the two controls; the kRAS promoter plasmids were not significantly different from the others (kRAS-500: 0.69 ± 0.11 and kRAS-324: 0.94 ± 0.16 fold RLU). HEK-293 cells were

transfected with each of the plasmids and treated with either 25 μM TMPyP2 or TMPyP4 for 48 hr (Figure 4-5C). TMPyP2 did not cause any significant effects with any of the four plasmids, and TMPyP4 did not affect EV (1.14 ± 0.21 fold RLU) or SV40 (1.01 ± 0.03 fold RLU) controls. Significant decreases in promoter activity were seen with TMPyP4 exposure to the kRAS-500 (0.69 ± 0.07 fold RLU) and kRAS-324 plasmids (0.61 ± 0.1 fold RLU). Because the G4-forming regions are contained within the kRAS-324 plasmid, this plasmid was chosen to pursue further studies as described above.

Site-directed mutagenesis allowed us to introduce G>T mutations within the near- or mid-G4 forming regions to ensure G4 formation was inhibited. We used the number of G>T mutations to minimize disruptions to transcription factor binding sites. Basal expression of these new plasmids was elucidated after 48 hr of transfection; no marked differences were noted amongst each of the plasmids (Figure 4-5D). When compared to kRAS-324 (0.94 ± 0.16), fold RLU of 324 mt Near was 1.31 ± 0.23, and 324 mt Mid was 0.96 ± 0.30. Like before, treatment with TMPyP2 did not bring about any significant changes to promoter activity with any of the plasmids, whereas both 25 and 50 μM TMPyP4 did (Figure 4-5E). The higher concentrations of TMPyP4 showed decreases in luciferase activity with both EV and SV40 plasmids. This dose-response may be due to the background effects of TMPyP4 by reducing luciferin glow due to the dark color. Thus, any dose-response related decrease of promoter activity was compared to the effects on non-G4 promoters by using a two-way ANOVA. Significant decreases ($p<0.05$) in promoter activity were observed in kRAS-324 upon 50 μM TMPyP4 treatment (26% decrease, 0.55 ± 0.09 fold RLU) and 324 mt Near (30% decrease, 0.51 ± 0.15 fold RLU) plasmids. 324 mt Mid showed a 21% increase in RLU, as compared to EV and SV40, and

it was 0.89 ± 0.14. Therefore, the crucial player for transcriptional silencing within the kRAS proximal promoter is most likely the $G4_{mid}$ higher order DNA structure, as promoter activity is suppressed when $G4_{mid}$ is stabilized with TMPyP4. Unlike $G4_{near}$, this new molecular structure is an ideal target for therapeutic development and rational drug design.

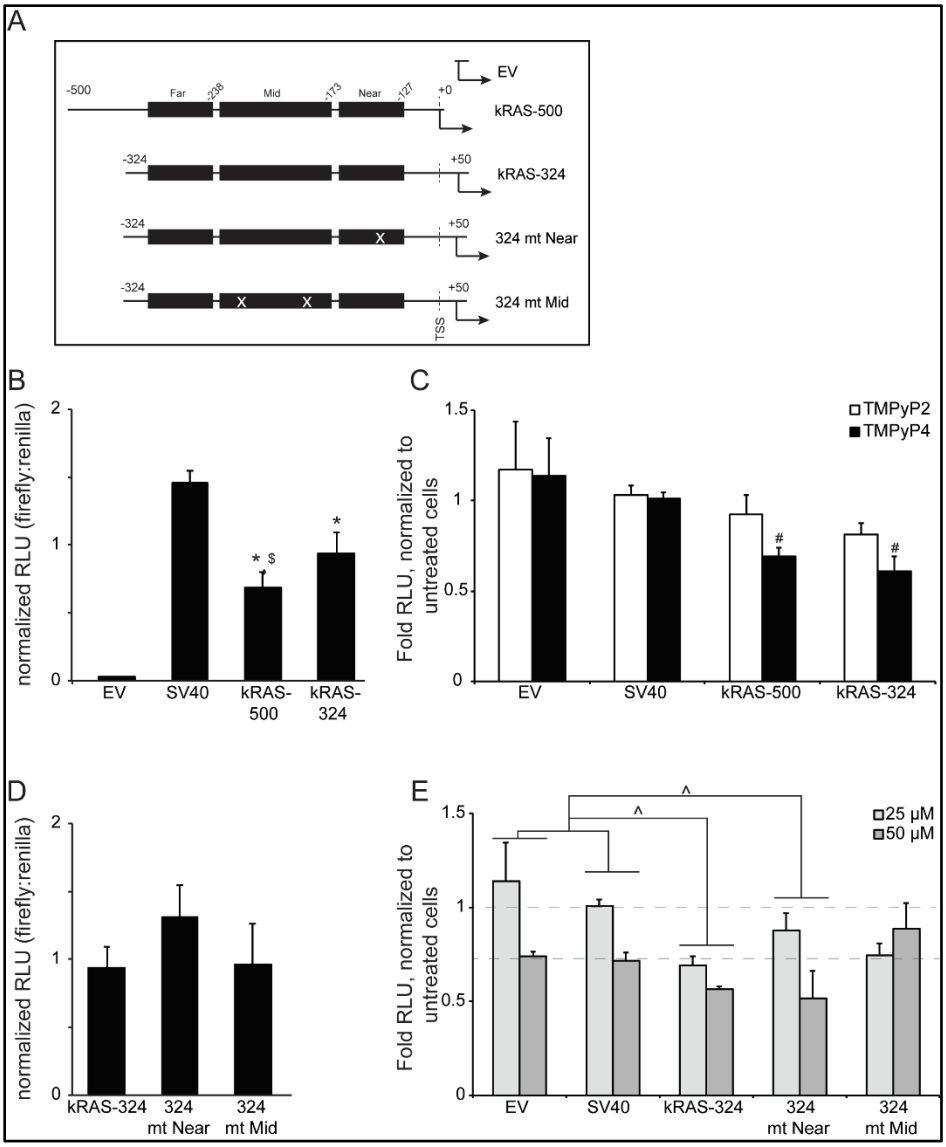


Figure 4-5. G4-mediated silencing is contained within the mid-region. Four plasmids were constructed using a promoterless empty vector (EV) pGL4.17 backbone to observe which regions within the kRAS promoter drive silencing activity (A) Basal expression of non-mutated plasmids showed no significance between each other but does compared to EV ($^{*}p<0.05$) and to SV40 ($^{\$}p<0.05$) by one-way ANOVA (B) Addition of 25 μM TMPyP2 and TMPyP4 to HEK-293 cells showed TMPyP4 being equally significant for each plasmid compared to untreated control ($^{\#}p<0.05$, as determined by one-way ANOVA) (C) Basal expression of kRAS-324 plasmid series harboring specific G4-knockout mutations for the near or mid regions showed no significant change in expression compared to wild-type (D) A dose response of TMPyP4 (25 μM and 50 μM) was examined in kRAS-324 and the mutated plasmids. By a two-way ANOVA, both kRAS-324 and 324 mt Near were significant in decreasing promoter activity ($^{\wedge}p<0.05$, as compared to EV and SV40). Because 324 mt Mid did not cause silencing, these observations suggest the G4-mediated silencing to be contained within the $G4_{mid}$ region (E) (Morgan et al. 2016)

4.3. Elucidation of $G4_{mid}$ in Physiological Conditions

With $G4_{mid}$ determined to be the biologically relevant structure in the above studies, we next looked to characterize G4 formation while mimicking the physiological conditions of a cell's nucleus. The intranuclear conditions are dehydrated water makes up approximately 60% of the solution mass along with extremely crowded with biomolecules (Cui et al. 2013). Biomolecules are important as they possess functions within living cells and inhabit 20-40% of the nucleus's volume. With this, we used a series of co-solvents to introduce these conditions to our G4 samples. The chosen co-solvents are neutral polymers polyethylene glycol (PEG), dextran sulfate, ficoll, and glycerol for molecular crowding and the corresponding monomers acetronitrile (MeCN), glucose, and sucrose as osmolytes. We incorporated both high and low mass polymers to increase net packing efficiency (Sharp 2016).

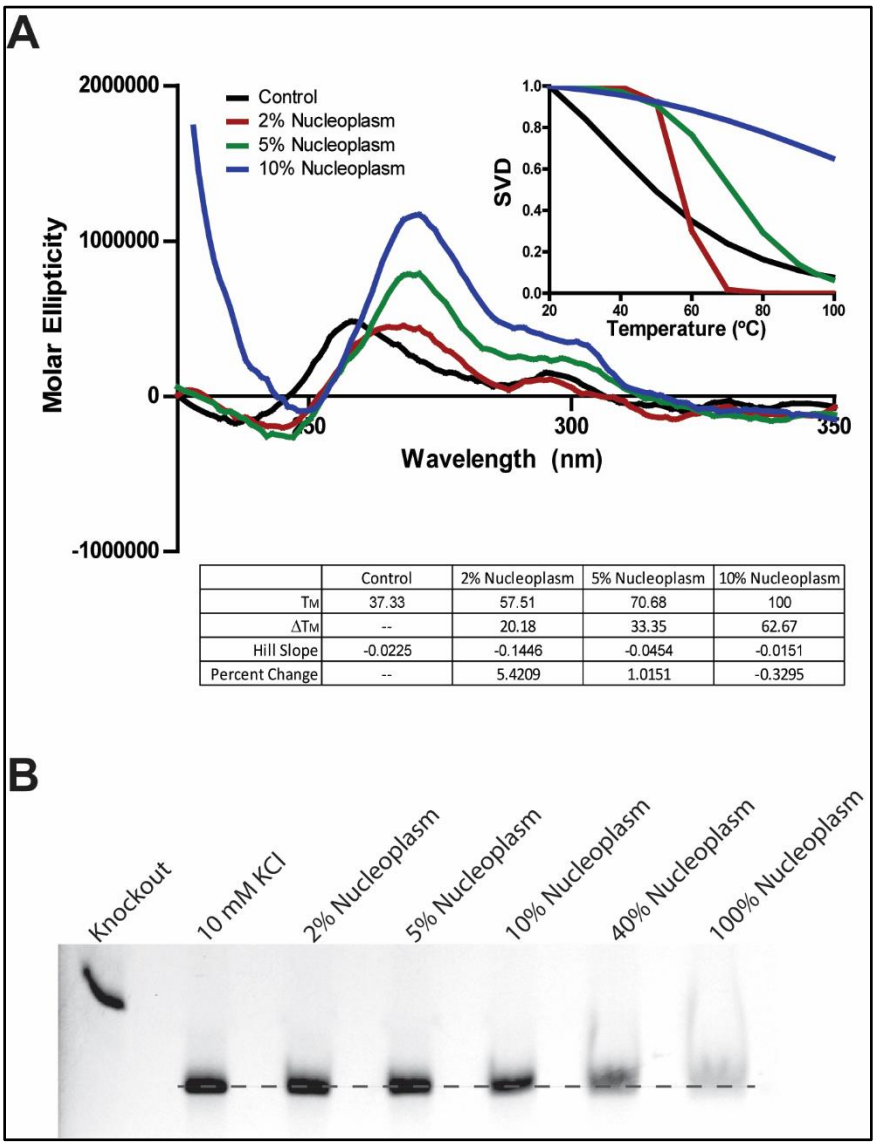
Along with these co-solvents, we utilized extracted nucleoplasm from MiaPaCa-2 cells to most likely mimic the nuclear conditions of a cell. We observed an increase in stability towards the mixed parallel/anti-parallel topology of $G4_{mid}$ with 10 mM KCl upon the addition of 2, 5, and 10% nucleoplasm (Figure 4-6A). Thermal stability increased from 37 °C to 57.5, 70.7, and 100 °C, respectively. With 2% nucleoplasm present, the number of isoforms decreased as the slope increased by 5.4% as compared to the 10 mM KCl control. Upon the addition of 5 or 10% nucleoplasm, negligible changes in slope were observed, compared to control. These lower concentrations (<10%) of extracted nucleoplasm were used in ECD studies, as 40 and 100% concentrations resulted in masking background on the Olis ECD instrument. The higher concentrations were used for EMSA studies where it was evident that there was a marked increase in the number

of isoforms present (Figure 4-6B). As compared to the linear knockout reference, $G4_{mid}$ forms intramolecular structures as noted by the downward migration pattern seen with KCl alone, as well as with each concentration of nucleoplasm; we observed an increase in the number of G4 isoforms existing as the amount of nucleoplasm increased. 2, 5, and 10% nucleoplasm each show crisp banding, similar to KCl alone, while 40 and 100% samples express more isoforms present with lighter and more dispersed banding patterns visible. Thus, we utilized lower concentrations, specifically 2 and 5%, of extracted nucleoplasm for the remainder of our studies.

The co-solvents were then characterized in order to determine the lowest concentrations (ranging from 0-40% at 10% increments) necessary for maximal G4 formation by $G4_{mid}$. By ECD spectral analysis we elucidated the saturation point in which this effect occurred for each co-solvent in the presence of 10 mM KCl (Figure 4-7). The saturating concentrations determined were 20% MeCN, 40% PEG, 20% ficoll, 20% glycerol, 10% glucose, 20% dextran sulfate, and 20% sucrose (Figure 4-8A). As compared to the 10 mM KCl control, these concentrations of co-solvents increased $G4_{mid}$ thermal stability by 24, 59, 17, 19, 20, 21, and 21 °C, respectively. A majority of co-solvents each markedly increased the slope, seen in the melting spectra in Figure 4-8A (MeCN by 141%, ficoll by 158%, dextran sulfate by 93%, and sucrose by 4,473%).

Next, we looked to compare and contrast the co-solvent ECD data with that of 2 and 5% extracted nucleoplasm (Figure 4-8B). In the case of 2% nucleoplasm, 20% MeCN closely resembled the same melting profile having no effect on thermal stability and increasing the slope, thus decreasing the number of $G4_{mid}$ isoforms, by 57%. With 5% nucleoplasm, other co-solvents held similar thermodynamic properties. With no effect on

thermal stability, 20% glycerol and 10% glucose had slight increasing effects on slope by 27 and 22%, respectively. Most of this ECD data is confirmed by EMSA in Figure 4-8C. Compared to linear control, knockout, $G4_{mid}$ forms intramolecular structures when co-solvents are present, with the exception of PEG, glucose, and sucrose. These three co-solvents resulted in intermolecular structure formation, as seen by the upward migration as compared to the KCl control. It is important to note that with inter- and intra-molecular G4 formation occurring, there is also banding coinciding with knockout migration indicating that the samples contained linear forms of the oligonucleotide as well. Collectively, the osmolyte MeCN and crowding agent glycerol seem to be promising co-solvents to move forward with in elucidating the predominant biologically active isoform of $G4_{mid}$.



	Control	2% Nucleoplasm	5% Nucleoplasm	10% Nucleoplasm
T_M	37.33	57.51	70.68	100
ΔT_M	--	20.18	33.35	62.67
Hill Slope	-0.0225	-0.1446	-0.0454	-0.0151
Percent Change	--	5.4209	1.0151	-0.3295

Figure 4-6. Using nucleoplasm as idealized solvent for physiological G4 formation. 2, 5, and 10% nucleoplasm increase $G4_{mid}$ formation and increase thermal stability. 2% nucleoplasm decreases the number of G4 isoforms present compared to control based on percent change of the Hill Slope (A) Electromobility shift assay describes intramolecular G4 formation with varying amounts of nucleoplasm. 2 and 5% nucleoplasm addition decreases the number of isoforms present compared to 40 and 100% (B).

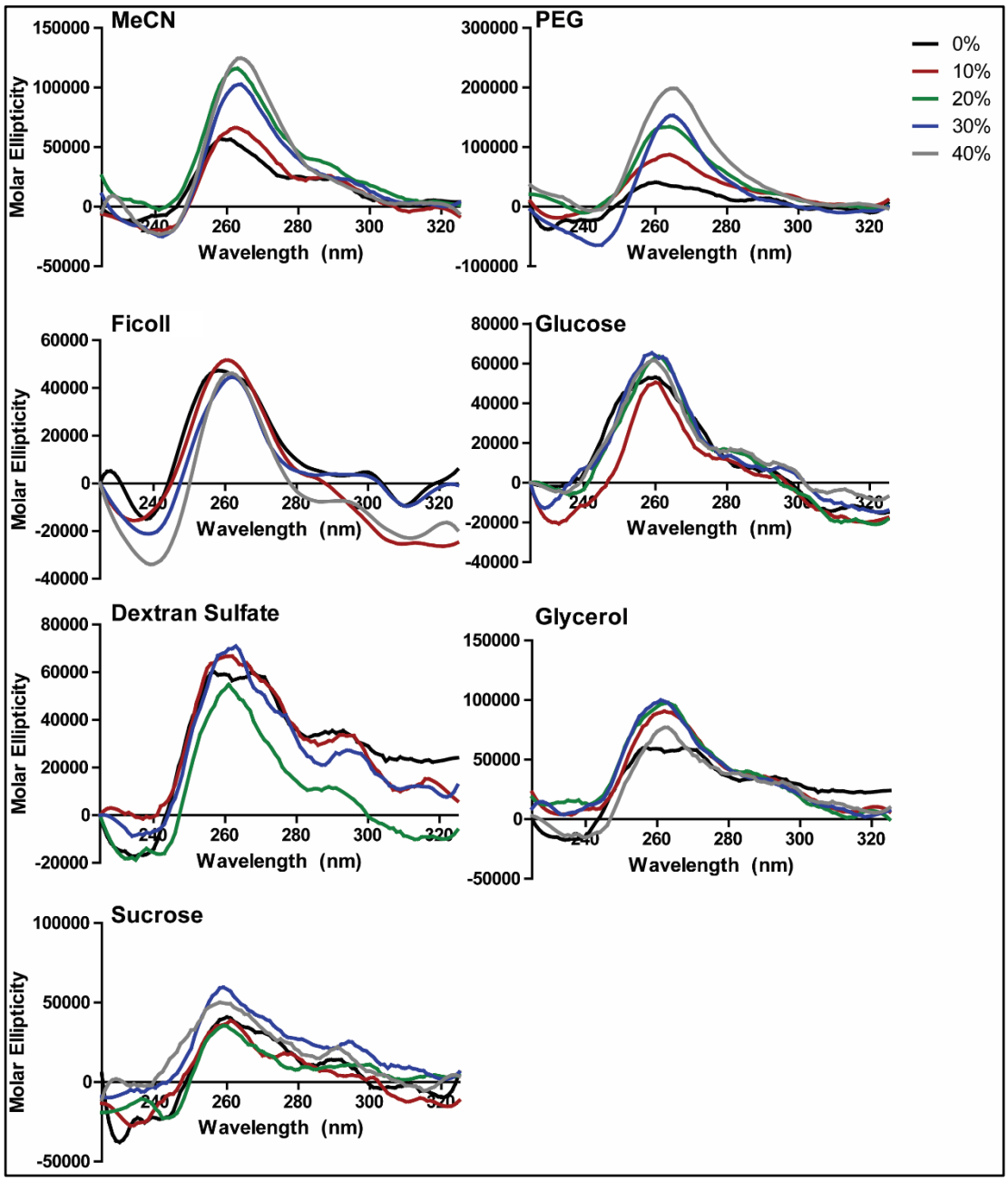


Figure 4-7. Determination of minimum concentration of co-solvent needed for maximum G4 formation. 0-40% gradients of various co-solvents, in 10% increments, were added to $kRAS_{mid}$. 20% MeCN, 40% PEG, 20% ficoll, 10% glucose, 20% dextran sulfate, 20% glycerol, and 20% sucrose allowed for maximal formation of $G4_{mid}$.

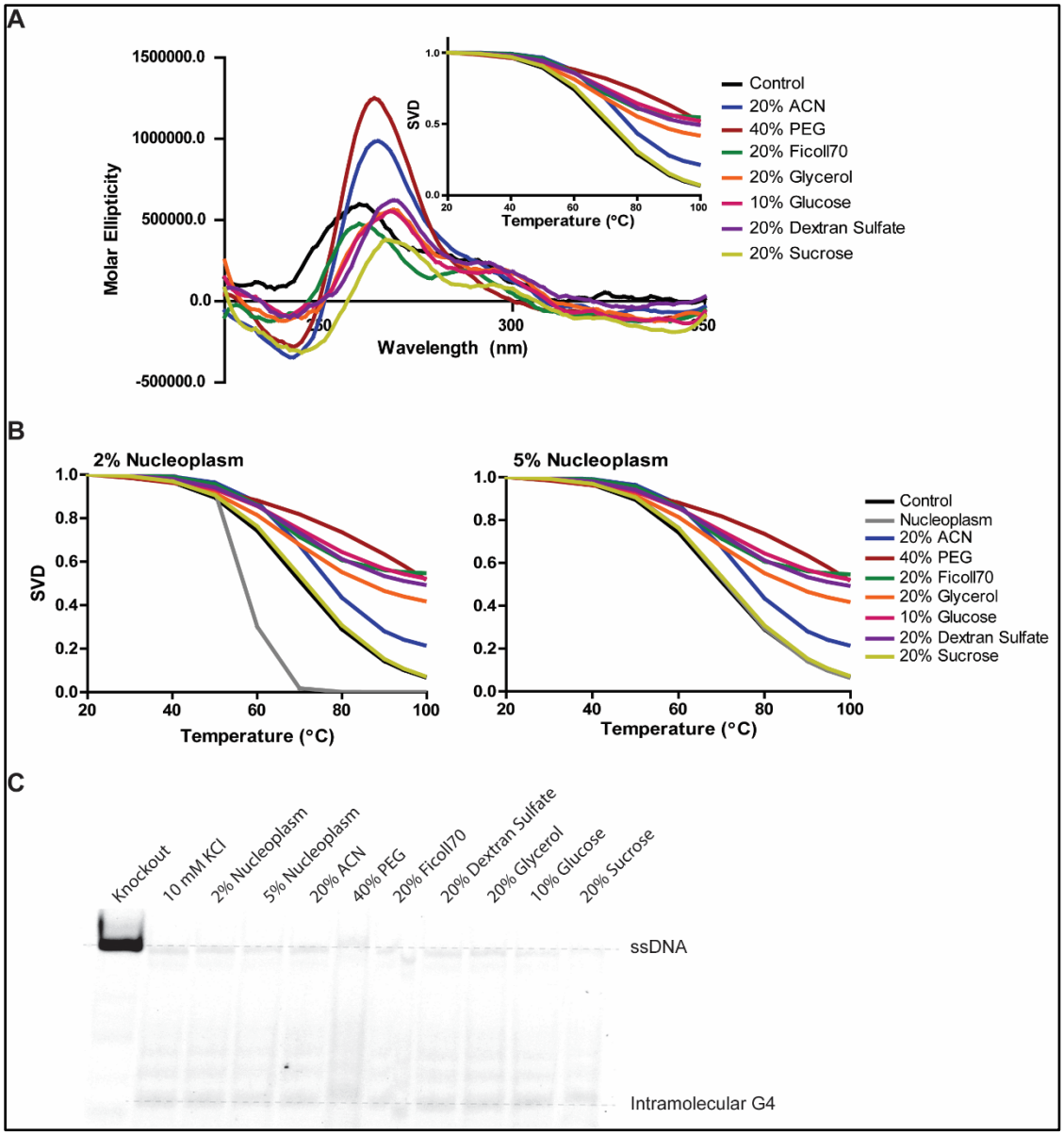


Figure 4-8. Comparing co-solvents to nucleoplasm. ECD spectra and melting profile of all co-solvents tested (A) 5% nucleoplasm and 20% sucrose share similar thermal profiles while 2% nucleoplasm has a markedly steeper slope than control or any co-solvent addition (B) EMSA shows decreased isoforms with nucleoplasm and MeCN compared to KCl. Most co-solvent additions led to intramolecular structure formation, but PEG, glucose, and sucrose caused intermolecular structures (C)

4.4. Targeting $G4_{mid}$ with Small Molecules

Characterizing the structure and average topology of the biologically active $G4_{mid}$ aids in G4 drug discovery efforts for kRAS. To identify small molecules to selectively stabilize $G4_{mid}$, a blind screen by FRET melt was performed for over 1,600 compounds from National Cancer Institute (NCI) Diversity Set III. Lead compounds from each 96-well plate were logged and were all repeated on the same plate to eliminate false positives (Figure 4-9). This assay utilizes a FRET probe that contains the $G4_{mid}$ region with a FAM-labeled fluorophore on the 5'-end and a TAMRA quencher on the 3'-end. G4 formation will bring the fluorophore and quencher together resulting in decreased fluorescent signal compared to the linear counterpart when the G4 does not form or unfolds. When screening small molecules, lead agents may have a decrease in initial fluorescence and will show an increase in melting temperature as they enhance thermodynamic stability of G4 formation.

FRET melt studies yielded 4 hit compounds that were moved forward for validation by ECD: NSC 317605, 106506, 274905, and 44750. ECD analysis was executed for both $G4_{mid}$ and $G4_{near}$ to understand the compound selectivity (Figure 4-10A). NSC 317605 did not change the number of isoforms for $G4_{near}$ or $G4_{mid}$; it moderately increased thermal stability of $G4_{near}$ by 9 °C and markedly enhanced $G4_{mid}$ stability by 18 °C. The compounds NSC 106506, 274905, and 44750 modestly destabilized $G4_{near}$ by 5, 3, and 4 °C, respectively, and pronouncedly affected $G4_{mid}$ with decreased melting temperatures of 10, 6, and 5 °C, respectively.

G4-Destabilizing — G4-Stabilizing

	1	2	3	4	5	6	7	8	9	10	11
A	-13.06	19.94	4.36	1.67	3.43	3.01	2.85	2.93	-0.73	0.88	-1.15
B	2.96	-6.24	3.47	-52.15	2.46	4.35	-52.15	-0.71	0.2	-3.61	0.73
C	2.59	18.31	-2.17	1.62	6.07	2.92	2.05	6.63	-0.98	0.67	0.39
D	2.69	35.22	1.62	2.02	7.82	-15.87	-3.07	0.55	18.82	-3.86	-3.67
E	40.92	-2.65	1.51	-0.41	0.61	6.75	-8.06	-52.15	-17.27	-0.84	0.05
F	8.08	-0.78	1.39	-4.69	-10.31	0.47	-1.04	-4.39	20.75	1.55	0.77
G	4.47	0.8	1.19	0.85	1.58	0.37	0.89	1.93	0.09	14.7	-0.5
H	1.74	0.17	1.33	4.73	2	1.44	1.31	0.38	0	-4.6	-1.26

Figure 4-9. FRET melt repeat of lead NCI compounds. Over 1,600 compounds from the NCI Diversity Set III were blindly screened by FRET melt. This 96-well plate shows the replicate of lead compounds suggesting approximately 7 hits based on melting temperature (red wells). Compounds exhibiting increased temperature changes greater than 5 °C are considered G4-stabilizing (red wells) agents while those decreasing melting temperature by greater than 5 C are G4-destabilizing (blue wells).

HEK-293 cells underwent a 48 hr transfection and treatment with 1 μM of each compound in EV and kRAS-324 plasmids. Though not all compounds affected G4 stability for either structure, each small molecule did significantly downregulate promoter activity in kRAS-324 (FL) compared to EV negative control and DMSO treated cells (Figure 4-10B). Compared to DMSO, the fold RLU showed NSC 106506 decreased promoter activity by almost 90% (0.089 ± 0.122), NSC 44750 by 73% (0.265 ± 0.078), NSC 274905 by 70% (0.299 ± 0.001), and NSC 317605 by 65% (0.345 ± 0.065). Collectively, from ECD and luciferase, the lead compound to pursue as a $G4_{mid}$ stabilizer and potential kRAS down-regulator further is NSC 317605, as it markedly reduces promoter activity and stabilizes $G4_{mid}$ greater than $G4_{near}$ (Figure 4-10C).

FRET melt was used to differentiate selectivity of 6 compounds over 3 G4s: $G4_{mid}$, $G4_{near}$, and MYC (Figure 4-11A). The test compounds included the cationic porphyrins, TMPyP2 and TMPyP4, with known properties for having no effect on stability and for stabilizing pan-G4s, respectively. These positional isomers act conversely as TMPyP2 does not impact G4 structures and TMPyP4 increases stability of all G4s non-selectively. The other 4 compounds tested share pharmacophore properties, such as a 4-ring aromatic structure and amine substitution. Also, NSC 176327 and 338258 are both ellipticines. The first report of quindoline i showed its ability to stabilize the MYC G4 while having anti-cancer activity in Hep G2 hepatocellular carcinoma cells (Ou et al. 2007). However, later it was understood that this inherent cytotoxicity was not dependent on the MYC G4 mechanism (Brown et al. 2011; Boddupally et al. 2012). Since then, quindoline i remains a G4 stabilizing compound effecting several structures and not limited to MYC with intracellular activities above G4s. NSC 176327 and 338258 were lead compounds

from other NCI Diversity Set screens that our lab identified to stabilize $G4_{near}$ and MYC G4s, respectively. NSC 338258 was reported as the first G4 selective agent with a validated mechanism of action dependent on G4 stabilization (Brown et al. 2011).

FRET melt studies with this series of compounds displayed different interactions depending on the G4 present. As predicted, TMPyP2 does not markedly change the thermal stability of each G4. Its isomer, as well as quindoline i, notably stabilized each G4 structure, as seen by the right-handed shift in thermal melt profiles (Figure 4-11B). In agreement with published studies, NSC 338258 stabilizes MYC and not $G4_{mid}$ or $G4_{near}$ (Brown et al. 2011). NSC 176327 did not affect $G4_{mid}$, but stabilizes both $G4_{near}$ and MYC. In contrast, our hit compound from the NCI Diversity Set III, NSC 317605, stabilizes $G4_{mid}$ and MYC, but not $G4_{near}$. From this cohort, NSC 176327 and 317605 remain interesting compounds in terms of kRAS G4 drug discovery as they each stabilize only one of the two G4s within the kRAS. After transfecting HEK-293 cells with kRAS-324 (FL) promoter plasmid, we noted that TMPyP2 and NSC 176327 do not markedly modulate promoter activity, but TMPyP4 and NSC 317605 significantly reduces promoter activity by 26% and 65%, respectively (Figure 4-11C). This further validates $G4_{mid}$ as the primary silencing structure within the kRAS promoter.

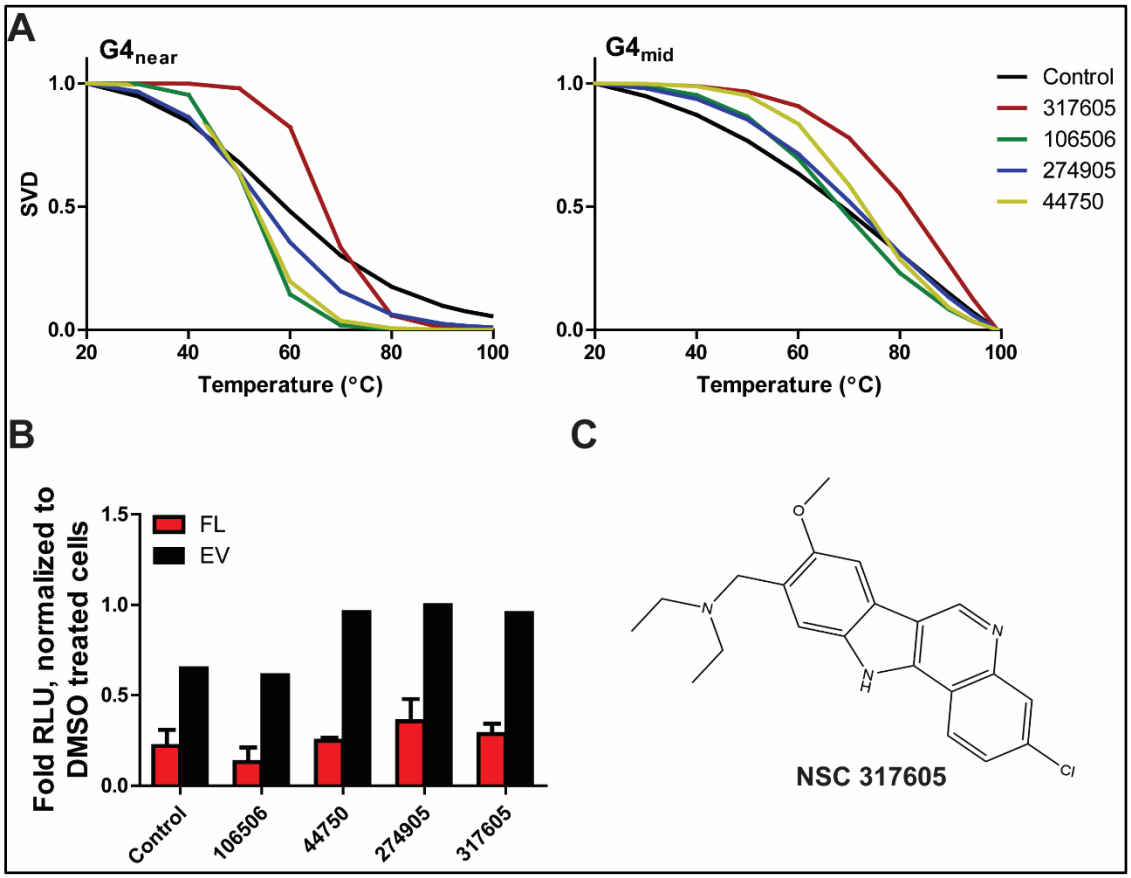


Figure 4-10. NCI lead compounds with $G4_{near}$ and $G4_{mid}$. Four lead compounds from the $G4_{mid}$ FRET melt screen were validated by ECD. NSC 317605 increased thermal stability greater than the other compounds for $G4_{mid}$ with little effect on $G4_{near}$ (A) All compounds decreased promoter activity in kRAS-324 (FL) as compared to EV (red bars are significant, $p<0.01$ as compared to DMSO treated cells, as determined by a two-tailed t-test) (B) Structure of NSC 317605 (C).

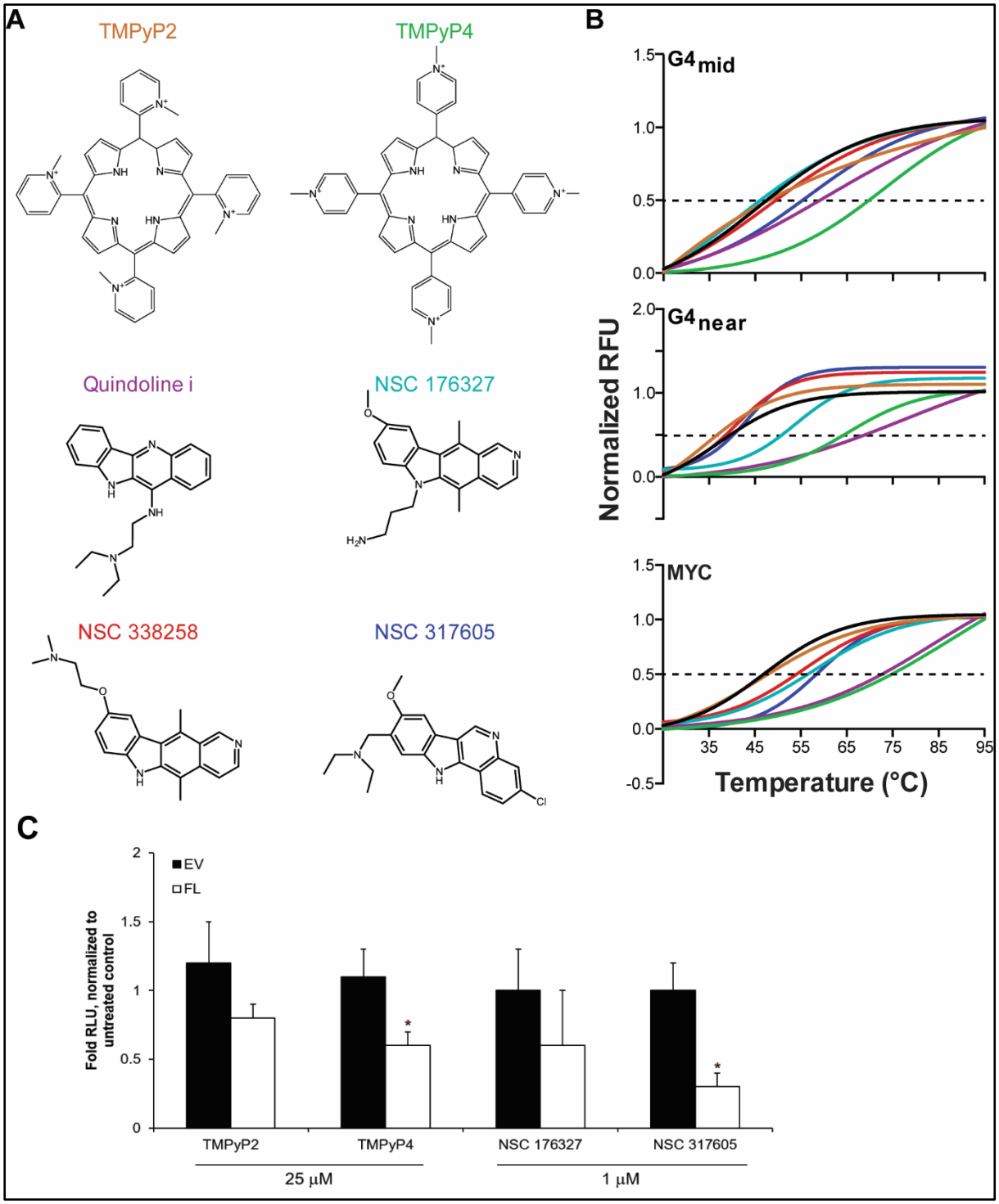


Figure 4-11. G4 stabilizing compounds and their varying selectivity. FRET melt of $G4_{mid}$, $G4_{near}$, and MYC with an array of G4-interactive compounds (A) show varying increases in stability (B) The universal stabilizing agents TMPyP4 and quindoline i, but not TMPyP2, broadly increase all G4 stability. NSC 176327 stabilizes $G4_{near}$ and MYC, not $G4_{mid}$, and NSC 317605 stabilizes $G4_{mid}$ and MYC, but not $G4_{near}$. G4-mediated silencing is seen in cells transfected with kRAS-324 and treated with 25 μM TMPyP4 or 1 μM NSC 317605, but not 25 μM TMPyP2 or 1 μM NSC 176327 ($^{*}p<0.05$ as compared to untreated control, as determined by a one-way ANOVA) (C).

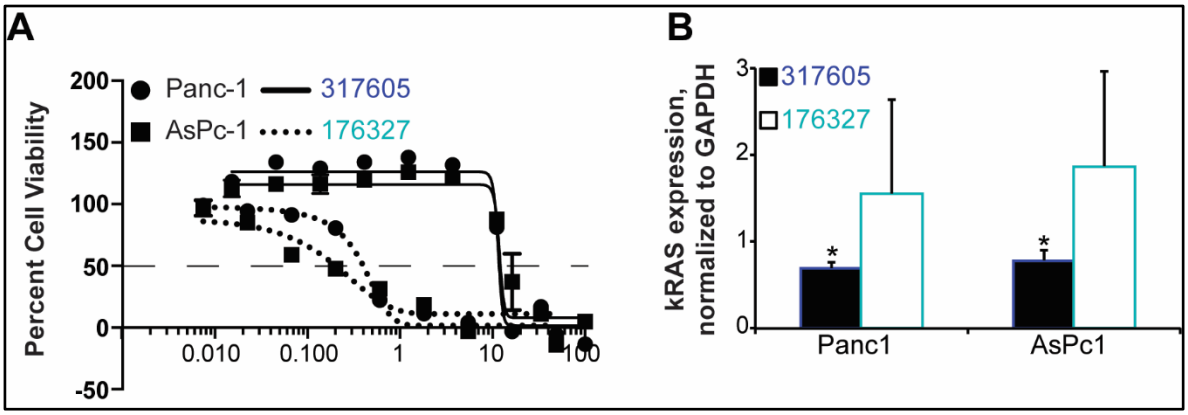


Figure 4-12. NSC 317605 and 176327 in pancreatic cancer cells. NSC 176327 is more cytotoxic than 317605 in Panc-1 and AsPc-1 cells (A) NSC 317605 decreased kRAS mRNA expression, unlike 176327 (*p<0.05, as compared to untreated control, determined by a one-way ANOVA) (B).

Next, we measured how each compound impacted cell viability amongst a pancreatic cancer cell panel (Figure 4-12A). Panc-1 cells are least addicted to kRAS mutation and AsPc-1 are an aggressive form of pancreatic cancer cells harboring a high addiction to kRAS mutation. After 24 hr exposure, NSC 317605 had IC_{50}s at 8.3 ± 0.6 and 11.72 ± 0.1, for Panc-1 and AsPc-1 respectively. NSC 176327 was more cytotoxic having 72 hr IC_{50}s of 0.4 ± 0.1 for both cell lines. Using these IC_{50}s to treat the cells for 24 hrs, qRT-PCR was performed to examine compound effect on kRAS mRNA expression, with respect to GAPDH (Figure 4-12B). Compared to untreated control cells, NSC 317605 significantly decreased kRAS mRNA levels while NSC 176327 had no effect. This further supports $G4_{mid}$ as the primary target for kRAS G4 drug discovery.

Simultaneously with the previous small molecule screening, we characterized compounds submitted by our collaborator Dr. Khondaker Miraz Rahman from King's College in London. We were searching for compounds selective for G4 structures within the kRAS promoter, specifically the mid region. The class of triaryl benzofurans examined consisted of 23 analogs. Thermodynamic stability was examined for $G4_{mid}$ and $G4_{near}$ upon compound addition (Figure 4-13A). Several compounds markedly increased $G4_{mid}$ stability over $G4_{near}$ such as KN-119, -159, -212, - 239,-257, and -267 while others acted on both G4s (KN-154, -217, -232, -237, -242, -247, and -272. Along with influencing thermodynamic stability, many of these compounds decreased promoter activity upon 48 hr treatment in kRAS-324 transfected HEK-293 cells (Figure 4-13B). These results indicate that compounds KN-159, -212, -237, -239, -242, and -272 should move into mechanistic confirmation studies using the pancreatic cancer cell line panel mentioned above.

Of these six compounds, most did not exhibit cellular cytotoxicity after 72 hr exposure in a panel of pancreatic cancer cells at concentrations up to 100 μM. KN-242 and KN-272 each had a dose-dependent cytotoxic effect on pancreatic cancer cells. KN-242 had IC_{50} values of 2 and 11 μM in Panc-1 and MiaPaCa-2 cells, respectively (cell viability with KN-242 treatment performed by Dr. Tracy Brooks, Figure 4-14B), and KN-272 had IC_{50} values of 2.1, 60, 6, and >100 μM in BxPc-3, Panc-1, MiaPaCa-2, and Capan-1 pancreatic cancer cells, respectively. Observation of these compounds with ECD showed that in the case of $G4_{mid}$, the structure topology shifted from predominantly parallel to more anti-parallel isoforms (Figure 4-14A). No change in loop directionality was examined for $G4_{near}$. This phenomenon is believed to be caused by the compound (structure shown in Figure 4-14D) interacting with the nucleotides contained within the G4 loops as it stacks above and/or below the structure. Similarly, KN-272 changes the topology of $G4_{mid}$ (Figure 4-15A) and not $G4_{near}$ (data not shown). Next, we determined each compounds' influence on kRAS mRNA expression. KN-242 significantly reduced kRAS expression, as well as the expression of other important genes involved in tumorigenesis, such as c-myc, ADAM-15, and Bcl-2 (qRT-PCR performed by Harshul Batra, Figure 4-14C)). Likewise, KN-272 had no marked impact on kRAS mRNA expression in wild-type kRAS cells, BxPc-3, or in kRAS mutated cells, MiaPaCa-2, compared to untreated cells (Figure 4-15B). (Figure 4-13 and 4-15A). This data suggests that KN-242, not KN-272, should undergo further studies investigating the biologically active kRAS G4.

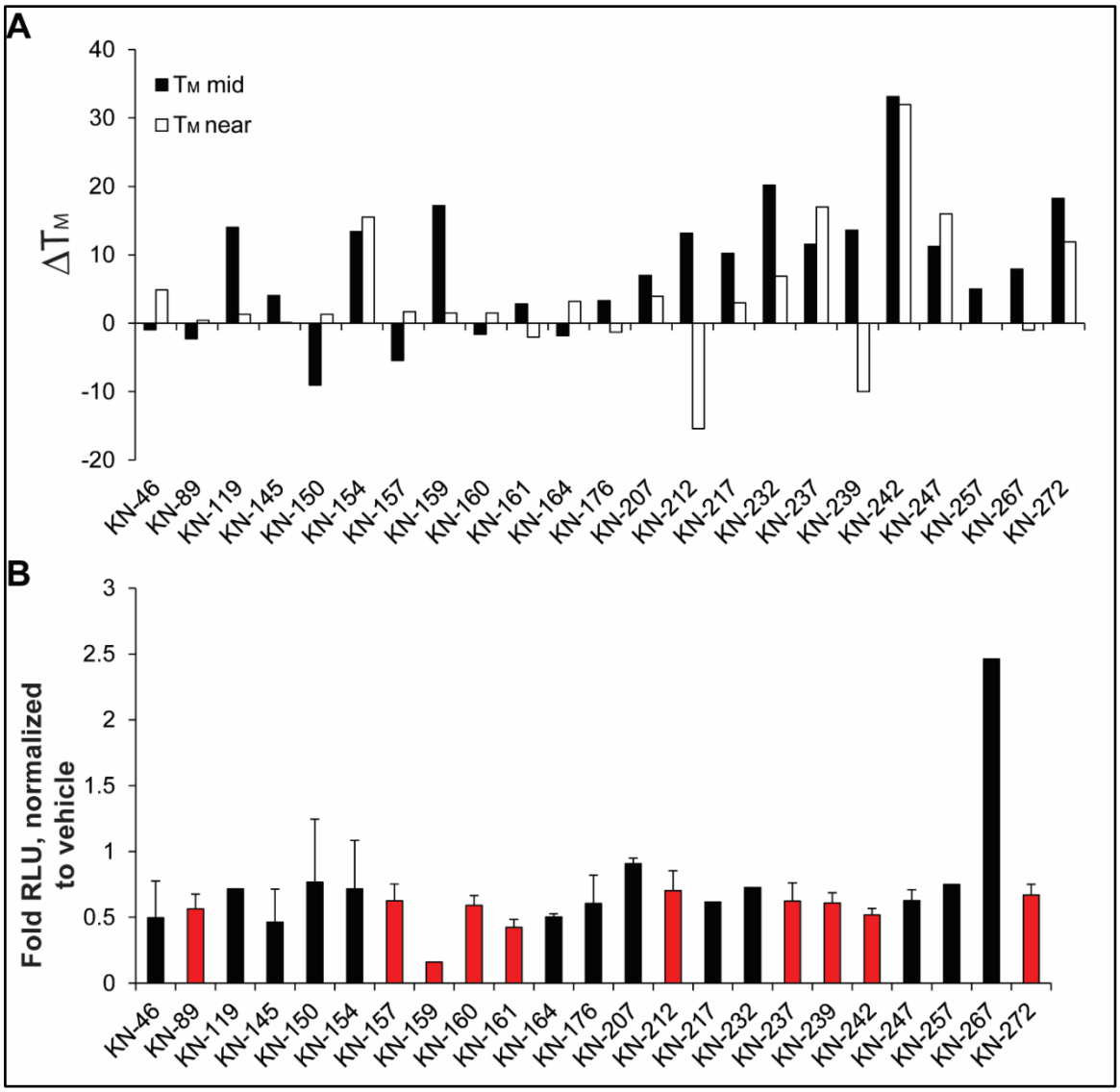


Figure 4-13. Triaryl benzofuran compound effect on G4 thermal stability and kRAS promoter activity. 23 compounds were screened by ECD to examine changes in thermal stability of $G4_{mid}$ and $G4_{near}$. 13 compounds markedly increased $G4_{mid}$ melting temperature compared to control while 7 increased that of $G4_{near}$ (A) 10 compounds significantly (red bars, determined by ANOVA with Dunnett post hoc analysis) decreased kRAS-324 promoter activity, 6 of which stabilize $G4_{mid}$ only and 3 that stabilize both $G4_{mid}$ and $G4_{near}$ (B).

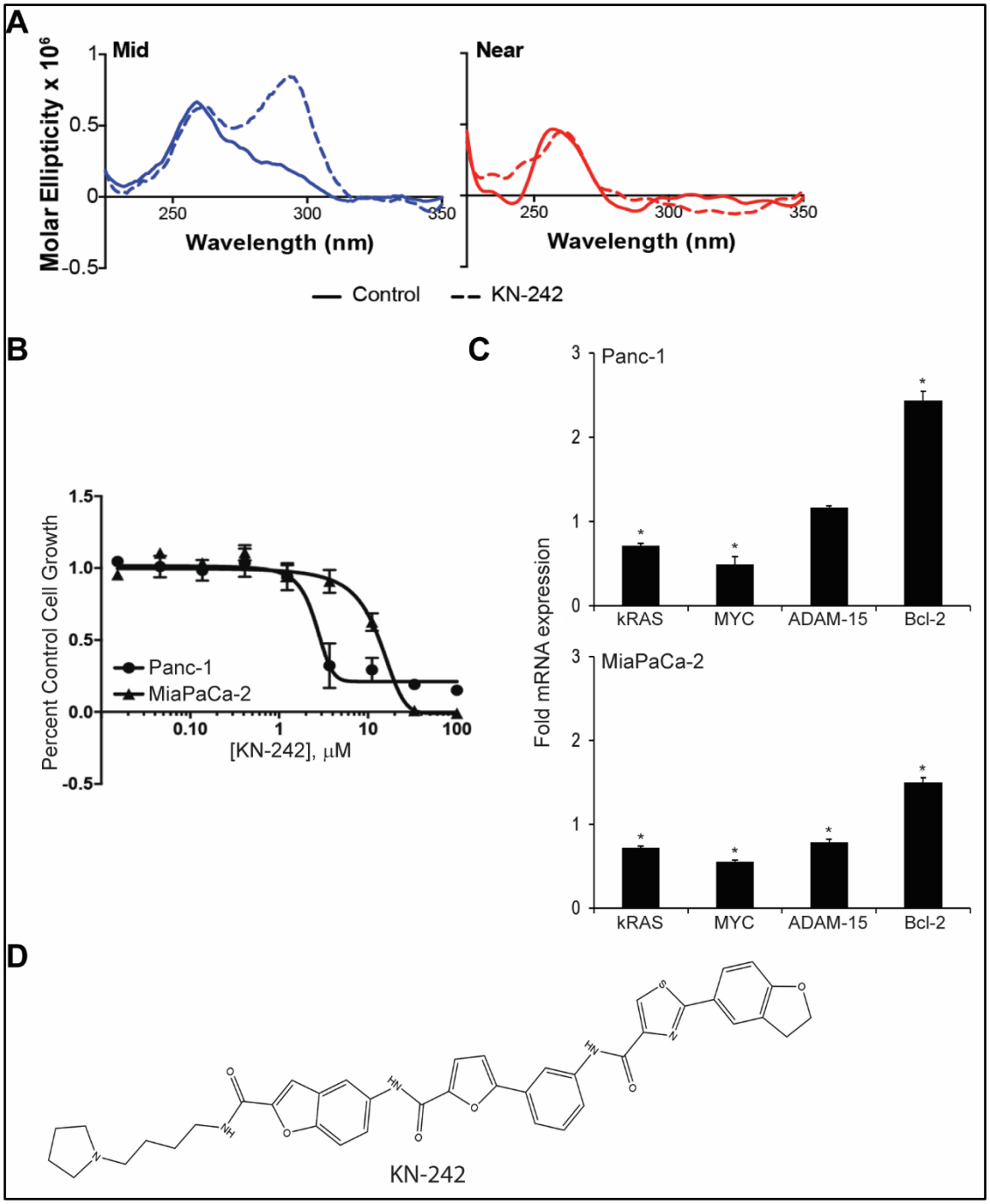


Figure 4-14. $G4_{mid}$ topology changes with KN-242. ECD spectra of $G4_{mid}$ changing from a predominantly parallel topology in control (solid lines) to a predominantly anti-parallel/mixed conformation upon KN-242 addition (dashed lines). No change in topology observed in $G4_{near}$ (A) KN-242 is more cytotoxic in the mildly addicted to kRAS pancreatic cancer cell line, Panc-1, compared to the more addicted, MiaPaCa-2 (B) Significant decreases in mRNA expression was seen for kRAS, MYC, and Bcl-2 within both pancreatic cancer cell lines. ADAM-15 was also significantly decreased in MiaPaCa-2 cells (*$p<0.05$, as compared to untreated control, determined by 2-tailed student t-test) (C) Chemical structure of KN-242 (D)

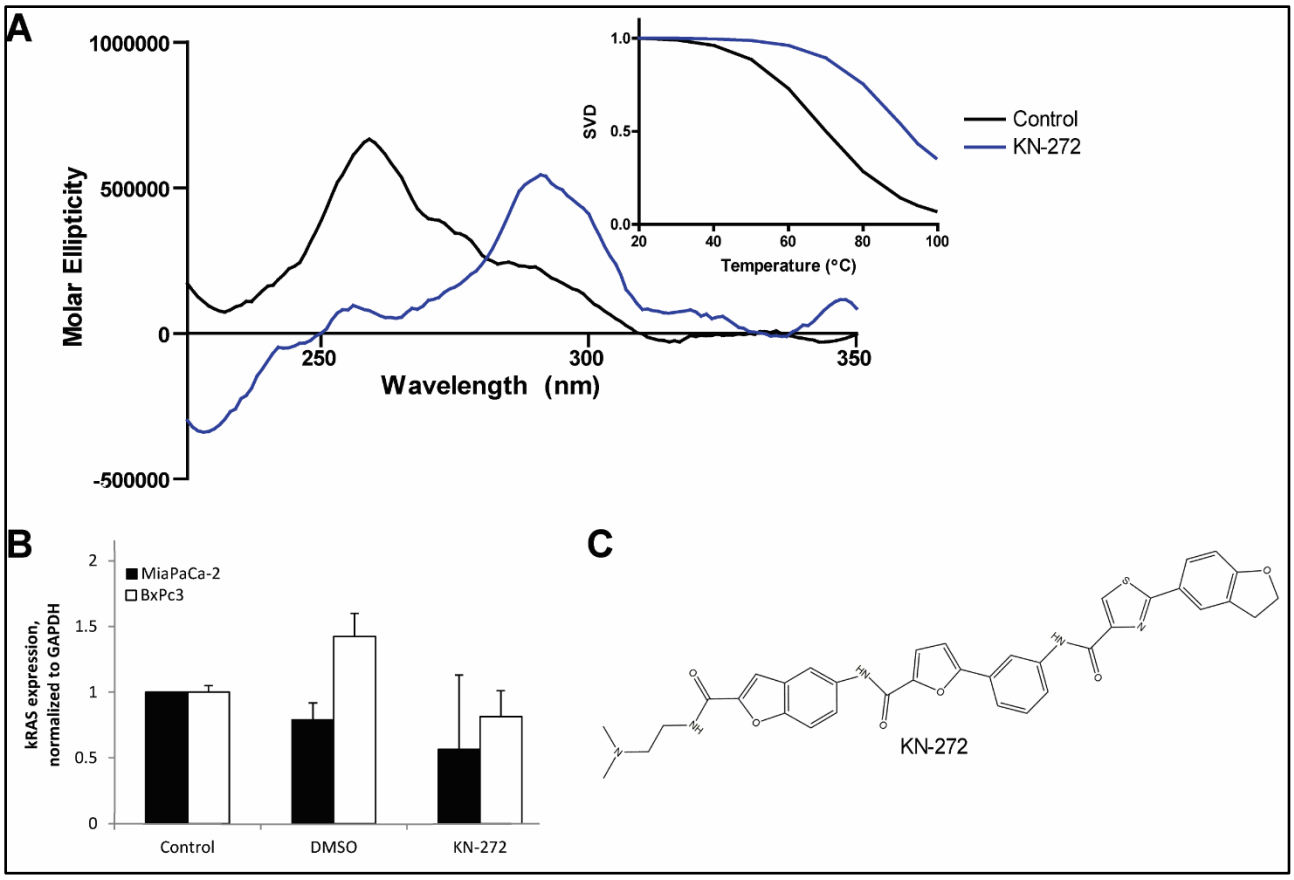


Figure 4-15. Changes in $G4_{mid}$ average topology and kRAS mRNA expression with KN-272. ECD spectra displays an exchange in topology for $G4_{mid}$ isoforms in the presence of KN-272 (blue lines) (A) No significant reduction in kRAS mRNA expression seen at the respective IC_{50}s in pancreatic cancer cells BxPc3 and MiaPaCa-2 (B) Chemical structure of KN-272 (C)

4.5. Discussion

Herein, we elucidated a new G4 in the kRAS core promoter by thoroughly examining the guanine-rich region that extends approximately 300 bases downstream from the transcriptional start site (TSS). Though three putative G4-forming regions seemed to exist within this area, only two formed stable inducible structures in different buffer solutions, those closest to the TSS. More significantly were our findings regarding $G4_{mid}$ being the biologically active structure within cells compared to what was previously described in literature as the function of what we termed $G4_{near}$. Preliminary characterization of $G4_{mid}$ suggests a predominant isoform being a tri-stacked mixed parallel/antiparallel structure with long loop lengths of up to ten nucleotides and the participating guanines are contained within the second, third, fifth, and sixth runs.

Extensive characterization work identified key co-solvents to be used while inducing $G4_{mid}$ formation by creating physiologically relevant conditions outside of a cell. We investigated the use of molecular crowding and dehydrating agents, alongside extracted nucleoplasm obtained from cells, in order to narrow the number of isoforms existing. This series of experiments suggests that 20% MeCN and 20% sucrose act similarly on G4 formation compared to extracted nucleoplasm and decrease the number of predominating isoforms. EMSA confirmed the presence of 20% MeCN and 2 or 5% nucleoplasm decreased the number of existing G4 isoforms, favoring parallel intramolecular structures.

The proximal promoter of kRAS is exceeding guanine rich (~75% G/C) and consists of the three described G4-forming regions within two nuclease hypersensitivity elements (Jordano & Perucho 1988). The region +/- 50 bp both upstream and

downstream of the TSS is responsible for transcription initiation while elements further upstream, such as the G4-forming regions, can act as silencing elements. Mutation of such elements has been reported to change basal promoter activity, but this was not observed with mutated $G4_{near}$ or $G4_{mid}$ which coincides with the knowledge of kRAS promoter activity being initiated closer to the TSS (Seenisamy et al. 2004; Wang et al. 2010).

Interestingly, our work describes a lack of silencing potential within the near G4-forming sequence, which goes against previous findings. It was reported that when a co-transfection of the human kRAS $G4_{near}$ sequence with a mouse kRAS promoter driven chloramphenicol acetyl transferase (CAT) plasmid decreased luciferase expression by 20% upon TMPyP4 treatment, compared to untreated control (Cogoi & Xodo 2006). The human and mouse sequences have an approximate 70% homology similarity (Yamamoto & Perucho 1988). Like wild-type, a mutant knockout mouse sequence of $G4_{near}$, increased basal expression indicating a silencing effect. This differs from the current study as we focused entirely on the human core promoter sequence of kRAS, using the entire core promoter with both wild-type and mutant knockouts. Thus, our work is more directly relatable to the endogenous state and depicting the biological function of $G4_{near}$ in humans; this was not entirely examined previously. Though the findings are dissimilar, our conclusion of G4-mediated silencing being contained within the mid-region does not go against previous reports of G4s within the kRAS core promoter. For example, within the first 500 bases downstream of the TSS, the addition of universal G4-stabilizing agent, TMPyP4, was reported to have decreased kRAS promoter activity, which we confirmed (Lavrado et al. 2015). The only difference is that we discovered that the silencing effect

is contained within the first 324 bases from the TSS, and narrowed it to the mid G4 region (-174 to -226 from the TSS).

Many studies have been performed looking at the regulation of the overall G4-forming region within the kRAS promoter. This has been performed through the use of protein transfection and G4-decoys that differ from G4s structurally based on locked or twisting base modifications. In each of these instances, G4 formation led to the transcriptional silencing previously described, but our findings indicate that this event is mediated by the formation of $G4_{mid}$ (Cogoi et al. 2008; Paramasivam et al. 2009; Cogoi et al. 2010). Notably, it is difficult to perform transcriptional regulation studies for kRAS as physical mapping of regulatory binding proteins has not been reported (Jordano & Perucho 1988; Yamamoto & Perucho 1988). However, through the use of PROMO virtual laboratory and Qiagen Champion ChiP Transcription Factor search portal, there are many potential transcription factor binding sites (consensus sequences) that reside within the kRAS core promoter, specifically in the G4-forming regions. There are several Sp1 sites surrounding the TSS and contained within the near and mid G4-forming regions, as well as sites for MAZ, E2F-1, AP-2, p53, and WT1. Knowing this, when considering the current studies using G>T mutations made within the near and mid regions, it can be assumed that there was disruption of transcription factor binding. These disruptions may have aided in the transcriptional silencing we observed, but our work co-transfecting wild-type or mutant human kRAS promoter sequence plasmids and TMPyP4 suggests that the G4-mediated silencing is more likely to occur due to $G4_{mid}$ than $G4_{near}$.

Knowing that $G4_{mid}$ is responsible for the G4-mediated silencing effect on kRAS expression, an initial search for G4-stabilizing small molecules was pursued. Utilizing NCI

Diversity Set III, the overall rate for positive leads was approximately 0.25% which further decreased to 0.0625% upon thermodynamic stability validation using ECD. Compound NSC 317605 markedly stabilized $G4_{mid}$ with a lesser effect on $G4_{near}$, but its selectivity is small, as it stabilizes other G4s such as MYC. Using NSC 317605 and the $G4_{near}$ stabilizing agent, NSC 176327, our previous findings of the silencing function being mediated by $G4_{mid}$ were confirmed by decreases in promoter activity and kRAS expression in cells. Several other compounds were screened as well but did not yield great success. Future studies will involve synthesizing analogues of NSC 317605 and investigating the role on $G4_{mid}$ stabilization and modulating kRAS transcriptional activity as performed here. Also, binding affinity and dissociation constants will be examined using competition dialysis and SPR, respectively. Similar to the current works, these analogues will be screened against several other known G4s like MYC, VEGF, Bcl-2, c-kit, hTERT, and ADAM-15 (Morgan & Brooks 2016). Lastly, any pharmacophores identified will undergo in vivo xenograft models using mice and zebrafish to explore efficacy and toxicological profiles, respectively.

kRAS is considered a 'holy grail' in terms of discovering targeted cancer therapeutics. Despite its clear validation as a therapeutic target, efforts to date have not yet yielded a successful clinical candidate. Previous approaches have focused primarily on suppressing the signaling pathways downstream from the activated mutant kRAS protein, competing with endogenous GTP, or inhibiting localization to the inner membrane with farnesyltransferase inhibitors (Ghobrial & Adjei 2002; Adjei 2001; Cho & Lee 2002; Queneau et al. 2001; Rowinsky et al. 1999; End 1999). The current studies present a new molecular target for drugging kRAS that is markedly different from previous tactics. While

transcriptional down regulation of kRAS has been validated as a clinical approach using siRNA and antisense RNA, no molecular target exists to mediate this down regulation with a small molecule approach (Gray et al. 1993; Zhang et al. 1993; Aoki et al. 1997). $G4_{mid}$ within the kRAS core promoter bridges this gap in transcriptional or translational approaches to downregulate the oncogene. Transcriptional control via small molecule stabilization of non-canonical DNA structures, specifically resulting in down regulation of kRAS, has not been formerly done. G4-mediated silencing will prevent continuous cell growth mediated by the mutated RAS protein and result in apoptosis. This cell death would preferentially occur within cancer cells harboring mutated kRAS or in cases where wild-type kRAS is overexpressed, as they are addicted to this oncogene. kRAS is dysregulated in almost 30% of all cancers, predominating in pancreatic (>90%), colorectal (45%), and lung (35%). Cumulatively, this work will positively impact cancer patients with kRAS addicted tumors, and provide a molecular model for development of additional therapy options for cancers harboring mutant or upregulated kRAS.

CHAPTER 5. TARGETING BCL-2 FOR NEUROPROTECTION

5.1. Introduction

Parkinson's disease (PD) affects seven to ten million people worldwide with one million sufferers in the United States alone (Sissi et al. 2011). According to the National Parkinson Foundation, approximately 50,000-60,000 citizens of the United States are diagnosed each year while many cases remain undetected. After Alzheimer's, PD is the second most common neurodegenerative disease worldwide (van der Heide & Smidt 2013). It causes serious complications which are the 14th leading cause of death for the nation. PD cases are mainly idiopathic and occur with increased age. There is no current cure for this disease, and treatment options depend on the symptoms expressed by the patient. Common symptoms include tremors, rigidity, bradykinesia, and postural instability (van der Heide & Smidt 2013). Treatments are very expensive as medicines cost up to $2,500 per year for a single patient and surgery averages around $100,000 (Paramasivan et al. 2007). Though medication and surgical therapy are utilized, the main goal for treating PD is helping the patient live with the disease by focusing on lifestyle modifications. Drug discovery efforts focus on attenuation of neuron death even though the exact pathogenesis is unknown due to its complexity (Rangasamy et al. 2010). One theory that remains is a mitochondrial dysfunction that leads to apoptosis signaling being activated resulting in aberrant dopaminergic neuronal cell death (Burbulla et al. 2010).

In wild-type animals, cell death phenotype was assessed with overexpression and knockout of Bcl-2 genes (Chipuk et al. 2010). It was concluded that a Bcl-2-dependent mechanism was behind the cell death signaling with the presence of caspase cleavage. The Bcl-2 family of proteins consists of both anti-apoptotic and pro-apoptotic members (Hetts 1998). They either promote or inhibit apoptosis by acting as checkpoints prior to cytochrome c release by controlling outer mitochondrial membrane integrity. Once cytochrome c is released from the intermembrane space of the mitochondria into the cytosol and caspases have been activated, anti-apoptotic members of this family can no longer prevent programmed cell death from occurring as the balance of proteins has been shifted due to apoptotic stress. Though there is no genetic disposition leading to increased loss of dopaminergic neurons, a key option for PD drug design and development is to inhibit apoptosis with Bcl-2 being a crucial mediator.

Aberrant Bcl-2 signaling can occur due to targeted gene disruption of Bax leading to increased programmed cell death (Akhtar et al. 2004). Shifting the threshold from Bax to Bcl-2 has significant potential to halt neuronal degeneration (van der Heide & Smidt 2013). In transgenic mice overexpressing Bcl-2, select regions of the brain possessed neuroprotective effects as there was an excess of living neurons compared to the Bcl-2 knockout mice which suffered from degeneration of neurons vastly separated in the brain (Farlie et al. 1995; Jean-Claude Martinou et al. 1994; Michaelidis et al. 1996). So far, targeted delivery options have been experimented with using pathogenic viral vectors to transport Bcl-2 fixed plasmids within neuronal cells. However, this delivery is limited due to it being an unsafe dosing regimen to convert from mice to humans (Shacka & Roth 2005). Thus, a novel molecular target for PD therapeutics is needed.

The cell's protector protein, Bcl-2, was identified after its isolation from B-cell lymphoma which suppressed apoptosis and caused oncogenesis. The human Bcl-2 gene isolated had a t(14:18) chromosomal translocation that caused the gene to be associated with over-activation of lymphoma cells (Tsujimoto et al. 1985). Under normal conditions, this gene is regulated by several gene regulatory proteins and transcription factors. The gene control region has two promoters, P1 and P2, located 1300 bases apart (Youdim & Arraf 2004). P1 is the predominant promoter involved in transcription. Interestingly, this promoter region consists of several Sp1 binding sites. Sp1 binding typically occurs on a consensus sequence consisting of many guanine and cytosine bases. These G-rich regions allow for G4 formation to occur within the Bcl-2 promoter and potentially affect transcription initiation (Figure 5-1).

Stabilizing G4s in promoter regions can activate or suppress transcription initiation by blocking specific gene regulatory proteins, transcription factors, and RNA polymerase II. Within the Bcl-2 proximal promoter there exists four such structures. $G4_0$ overlaps the TSS thus blocking RNA pol II binding and causing transcriptional silencing. The biological function behind $G4_2$ stabilization has yet to be uncovered, but $G4_3$ was shown to act as a silencing agent. However, our G4 of interest is $G4_1$. The gene regulatory protein WT1, a product of Wilm's tumor suppressor gene, binds approximately 27 to 35 base pairs upstream of the predominant Bcl-2 promoter and negatively regulates transcription (Heckman et al. 1997). WT1 cannot bind to secondary DNA structures such as G4s as it can only bind to ssDNA. Thus, we hypothesize that stabilization of this structure with small molecules has may prevent WT1 binding and cause upregulation of Bcl-2 transcription resulting in neuroprotection. Also, we predict that selective binding and stabilization of the

Bcl-2 $G4_1$ by specific compounds, such as cannabinoid derivatives, will not only allow transcription initiation activation resulting in neuroprotection, but also, have the ability to readily cross the blood brain barrier allowing for the penetrance of neuronal cells. This unique approach to Parkinson's disease therapeutics was accomplished by the following aims.

Specific Aim 1. Elucidate the predominant G4 stabilized with cannabinoid derivatives and assess transcription initiation potential.

Specific Aim 2. Demonstrate the in vitro neuroprotective potential of Bcl-2 G4-stabilizing compounds, and establish the mechanisms of action related to Bcl-2 in human dopaminergic neuronal cells.

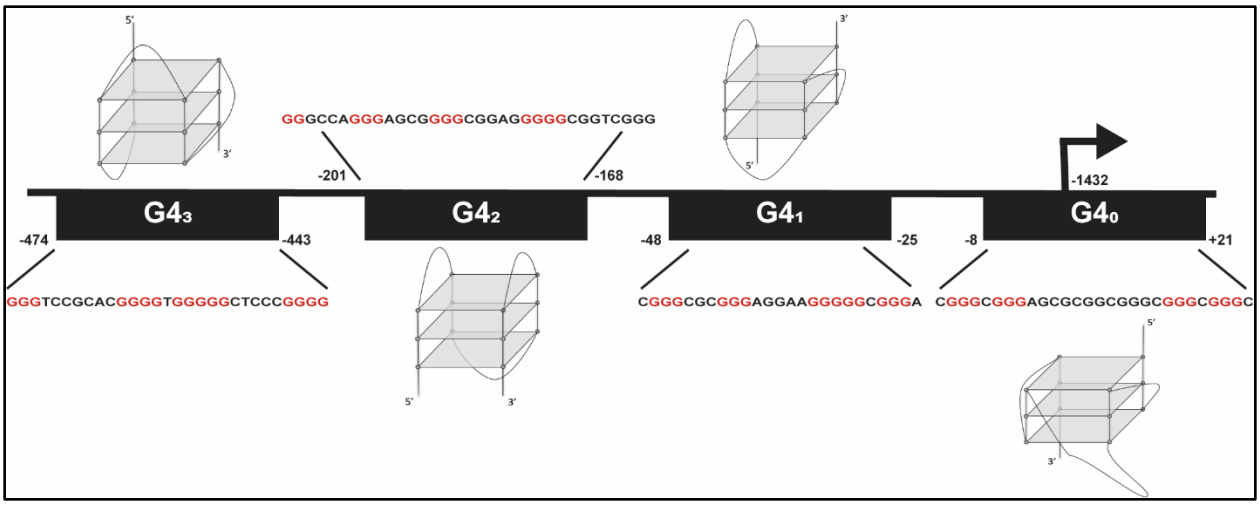


Figure 5-1. Complexity within the Bcl-2 promoter. There exist four G4s within the Bcl-2 proximal promoter region. $G4_0$ is a tri-stacked parallel structure located on the TSS having a hairpin forming the second loop that acts as a silencing element (Onel et al. 2016). $G4_1$ is immediately downstream from the TSS, acting as an enhancer, and forms tri-stacked mixed parallel/antiparallel structures (Dai et al. 2006).$G4_2$ is further downstream from the TSS forming a tri-stacked parallel structure with an unknown transcriptional mechanism (Onyshchenko et al. 2009). Lastly, $G4_3$ is the farthest from the TSS and forms a tri-stacked parallel structure that acts as a transcriptional silencer like $G4_0$ (Sun et al. 2014).

5.2. Exploring Physiologically Relevant Conditions for Small Molecule Targeting

The Bcl-2 promoter is excessively guanine dense with four G4-forming regions within 500 bp of the TSS. In 2006, NMR studies described a major intramolecular mixed parallel/antiparallel G4, Pu39, to exist from -1457 to -1480 (Dai et al. 2006). A few years later, with the use of invading short peptide nucleic acids (PNAs), G4 formation was observed in tandem to Pu39 (Onyshchenko et al. 2009). Termed bcl2G4-2, this G4-forming sequence is 51 bp and located between -1600 and -1633 and has a parallel average topology. Following this discovery, another mixed parallel/antiparallel structure, p32, was found to form -443 bp upstream from the TSS (Sun et al. 2014). More recently, a G4-forming region identified to overlap the TSS (Onel et al. 2016). P1G4 is a parallel structure harboring a hairpin loop. Due to the knowledge that these four major G4s can form within the Bcl-2 promoter, it suggests that there exists several isoforms of each of these structures as well.

Based on this complexity within the Bcl-2 promoter, we have labeled the G4s based on their proximity to the TSS. $G4_0$ is the most recent structure characterized with a hairpin loop, $G4_1$ is the first Bcl-2 G4 described, $G4_2$ was elucidated using PNAs, and, lastly, $G4_3$ is a mixed parallel/antiparallel structure furthest from the TSS (Figure 5-1) (Table 5-1). Based on the number of structures capable of forming, it is hard to predict which G4 may be the predominant structure *in vivo* as there are several other factors to consider physiologically that are hard to replicate *in vitro* like molecular crowding and dehydration.

Table 5-1. Bcl-2 oligonucleotide sequences.

$G4_0$	5’CGGGCGGGAGCGCGGCGGGCGGGCGGGC3’
$G4_1$	5’AGGGGCGGGCGCGGGAGGAAGGGGGCGGGAGCGGGGC3’
$G4_1$ FRET	5’[6~FAM]AGGGGCGGGCGCGGGAGGAAGGGGGCGGGAGCGG GGC[TAMRA~6~FAM]
$G4_2$	5’CGGGCCAGGGAGCGGGGCGGAGGGGGCGGTCGGGT3’
$G4_3$	5’GCTGGGGTCCGCGACGGGGTGGGGGCTCCCGGGGAAC3’

For the investigation of physiologically relevant conditions on Bcl-2 $G4_1$, we utilized extracted nucleoplasm and a variety of co-solvents. Compared to the 10 mM KCl control, the incorporation of extracted nucleoplasm (2, 5, and 10%) led to a significant reduction in the number of isoforms existing as the slopes increased by 17%, 1,292%, and 52%, respectively (Figure 5-2A). With this reduction, there was only an increase in G4 stability amongst the 10% nucleoplasm samples by roughly 11 °C. Next, minimal concentrations of co-solvents were examined that induce maximal G4 formation (spectra not shown). Using 40% MeCN, 10% ficoll, 30% PEG, 10% dextran sulfate, 10% sucrose, and 40% glucose, we observed an increase in slope upon the addition of dextran sulfate and sucrose indicating fewer G4 isoforms (Figure 5-2B). Of these co-solvents, MeCN and PEG act as stabilizing agents with 22 and 10 °C shifts, respectively. The remaining co-solvents had no marked effect on G4 stability as indicated by a temperature change of <5 °C as compared to control.

With this knowledge, we wanted to narrow the use of co-solvents to those comparable to extracted nucleoplasm (Figure 5-2C). Co-solvents ficoll, sucrose, and glucose closely relate to 2% nucleoplasm as there is no effect on thermal stability. However, sucrose is the only co-solvent with no effect on the number of isoforms existing as ficoll and glucose both increase G4 polymorphic nature with a reduction in slope by 17 and 52%, respectively. 5% nucleoplasm is similar to ficoll, dextran sulfate, sucrose, and glucose. Each of these decrease the slope resulting in more isoforms. Lastly, PEG is the only co-solvent closely related to 10% nucleoplasm with a 30% reduction in slope. MeCN acts as a stabilizing agent to $G4_1$ compared to any amount of nucleoplasm, as well as, the KCl control as seen in Figure 5-2B. Collectively, 5% nucleoplasm and its related co-

solvents should be used for further structural elucidation of the predominant physiologically relevant $G4_1$.

Narrowing down the number of G4 species to its physiologically relevant structure will aid in drug discovery efforts for $G4_1$. Similar to the MYC G4 studies previously described, we performed a small molecule screen with FRET melt under three conditions: control buffer, 10 mM KCl, and 10 mM KCl + 40% MeCN (Figure 5-3). The screen consisted of a wide variety of compounds ranging from known G4-stabilizing agents to natural product derivatives isolated or synthesized at the University of Mississippi. The same 96-well plate of compounds was used for screening under each of the three conditions. There was an 18% reduction in hit rate when comparing the control buffer (52%) to KCl + MeCN additions (30%). The lead compounds observed in the presence of co-solvent were next validated by ECD (Figure 5-4).

Of the six lead stabilizing compounds from FRET melt, two significantly increased $G4_1$ thermal stability as noted by a >5 °C melting temperature shift. Endocannabinoid derivatives EC-2014-01 and E-5-12-21 enhanced thermal stability by 6 and 15 °C, respectively, while the other compounds had no effect. Next, these compounds were tested against $G4_0$ and $G4_3$ as these structures possess a known biological role based on previous literature, both acting as silencing elements (Sun et al. 2014; Onel et al. 2016). $G4_0$ was destabilized upon addition of each compound except KN-242 which yielded no effect on G4 stability. Notably, compounds EC-2015-04, EC-2014-09, and E-5-12-21 reduced thermal stability by 9, 9, and 11 °C, respectively. $G4_3$ was not influenced by EC-2014-09 or KN-242, but EC-2014-01 reduced stability by 11 °C while the other three compounds prevented the induction of structure formation altogether. Thus,

compounds EC-2014-01 and E-5-12-21 act selectively on $G4_1$ over its neighbors, making them lead compounds for cellular studies to confirm the mechanism of action to be G4-dependent.

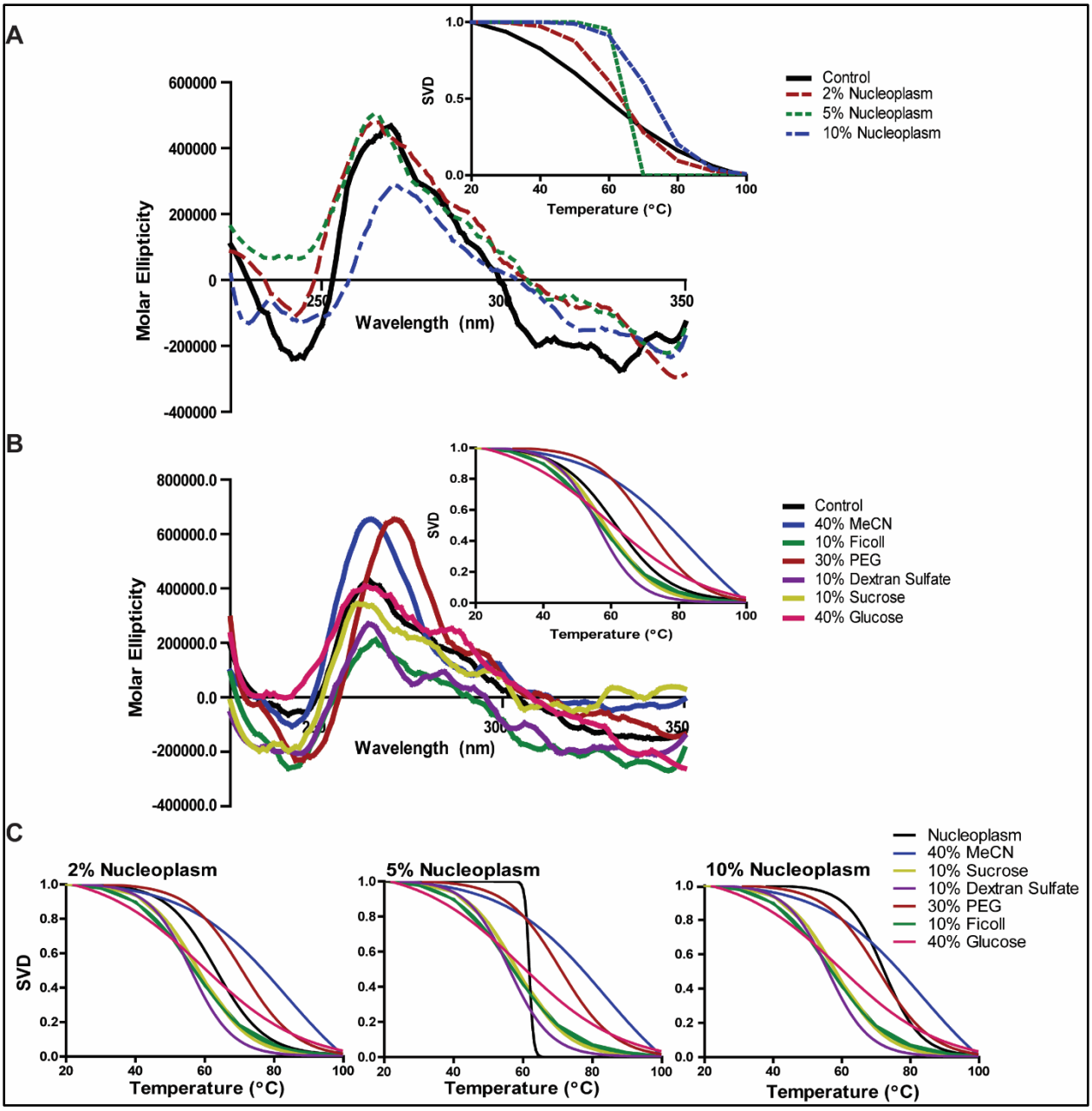
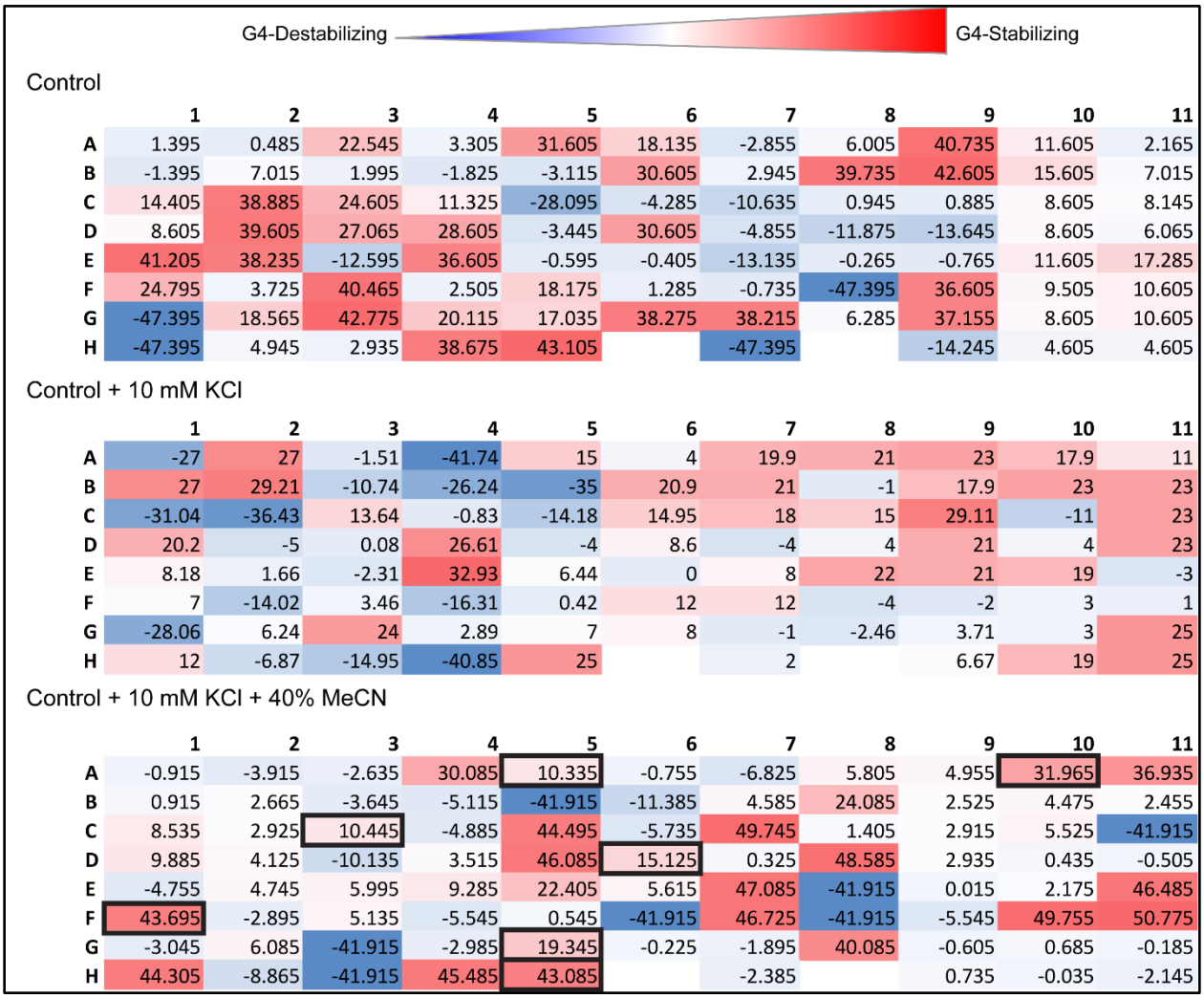


Figure 5-2. Extracted nucleoplasm and other co-solutes reduces average topology through physiological mimicry. Similar to 10 mM KCl control, 2 and 5% nucleoplasm induced stable G4 formation with both decreasing the number of isoforms without increasing G4 thermal stability. 10% nucleoplasm decreased the amount of stable, inducible G4 formation and increased thermal stability (A) A wide variety of co-solvents were utilized to observe any effect on G4 formation compared to 10 mM KCl control. MeCN markedly increases G4 formation and thermal stability. PEG, like MeCN, increases thermal stability, but shifts G4 formation more toward ssDNA by a maxima around 270 nm. The other so-solvents closely resemble control (B) When comparing co-solvents to nucleoplasm, most co-solutes mimic 2% nucleoplasm (C)



Control

	1	2	3	4	5	6	7	8	9	10	11
A	1.395	0.485	22.545	3.305	31.605	18.135	-2.855	6.005	40.735	11.605	2.165
B	-1.395	7.015	1.995	-1.825	-3.115	30.605	2.945	39.735	42.605	15.605	7.015
C	14.405	38.885	24.605	11.325	-28.095	-4.285	-10.635	0.945	0.885	8.605	8.145
D	8.605	39.605	27.065	28.605	-3.445	30.605	-4.855	-11.875	-13.645	8.605	6.065
E	41.205	38.235	-12.595	36.605	-0.595	-0.405	-13.135	-0.265	-0.765	11.605	17.285
F	24.795	3.725	40.465	2.505	18.175	1.285	-0.735	-47.395	36.605	9.505	10.605
G	-47.395	18.565	42.775	20.115	17.035	38.275	38.215	6.285	37.155	8.605	10.605
H	-47.395	4.945	2.935	38.675	43.105		-47.395		-14.245	4.605	4.605

Control + 10 mM KCl

	1	2	3	4	5	6	7	8	9	10	11
A	-27	27	-1.51	-41.74	15	4	19.9	21	23	17.9	11
B	27	29.21	-10.74	-26.24	-35	20.9	21	-1	17.9	23	23
C	-31.04	-36.43	13.64	-0.83	-14.18	14.95	18	15	29.11	-11	23
D	20.2	-5	0.08	26.61	-4	8.6	-4	4	21	4	23
E	8.18	1.66	-2.31	32.93	6.44	0	8	22	21	19	-3
F	7	-14.02	3.46	-16.31	0.42	12	12	-4	-2	3	1
G	-28.06	6.24	24	2.89	7	8	-1	-2.46	3.71	3	25
H	12	-6.87	-14.95	-40.85	25		2		6.67	19	25

Control + 10 mM KCl + 40% MeCN

	1	2	3	4	5	6	7	8	9	10	11
A	-0.915	-3.915	-2.635	30.085	10.335	-0.755	-6.825	5.805	4.955	31.965	36.935
B	0.915	2.665	-3.645	-5.115	-41.915	-11.385	4.585	24.085	2.525	4.475	2.455
C	8.535	2.925	10.445	-4.885	44.495	-5.735	49.745	1.405	2.915	5.525	-41.915
D	9.885	4.125	-10.135	3.515	46.085	15.125	0.325	48.585	2.935	0.435	-0.505
E	-4.755	4.745	5.995	9.285	22.405	5.615	47.085	-41.915	0.015	2.175	46.485
F	43.695	-2.895	5.135	-5.545	0.545	-41.915	46.725	-41.915	-5.545	49.755	50.775
G	-3.045	6.085	-41.915	-2.985	19.345	-0.225	-1.895	40.085	-0.605	0.685	-0.185
H	44.305	-8.865	-41.915	45.485	43.085		-2.385		0.735	-0.035	-2.145

Figure 5-3. FRET melt drug discovery application of co-solvent MeCN. The same 96-well plate was screened in control, KCl, and KCl+MeCN conditions to examine any differential hits. The screening rate for positive compounds decreased from 52% in control to 30% upon KCl+MeCN addition. G4-stabilizing compounds are observed in red while blue wells indicate G4-destabilizing agents. White wells depict compounds with no effect on G4 stability. G4-stabilizing small molecules persistent amongst each of the three screening conditions are outlined in black boxes.

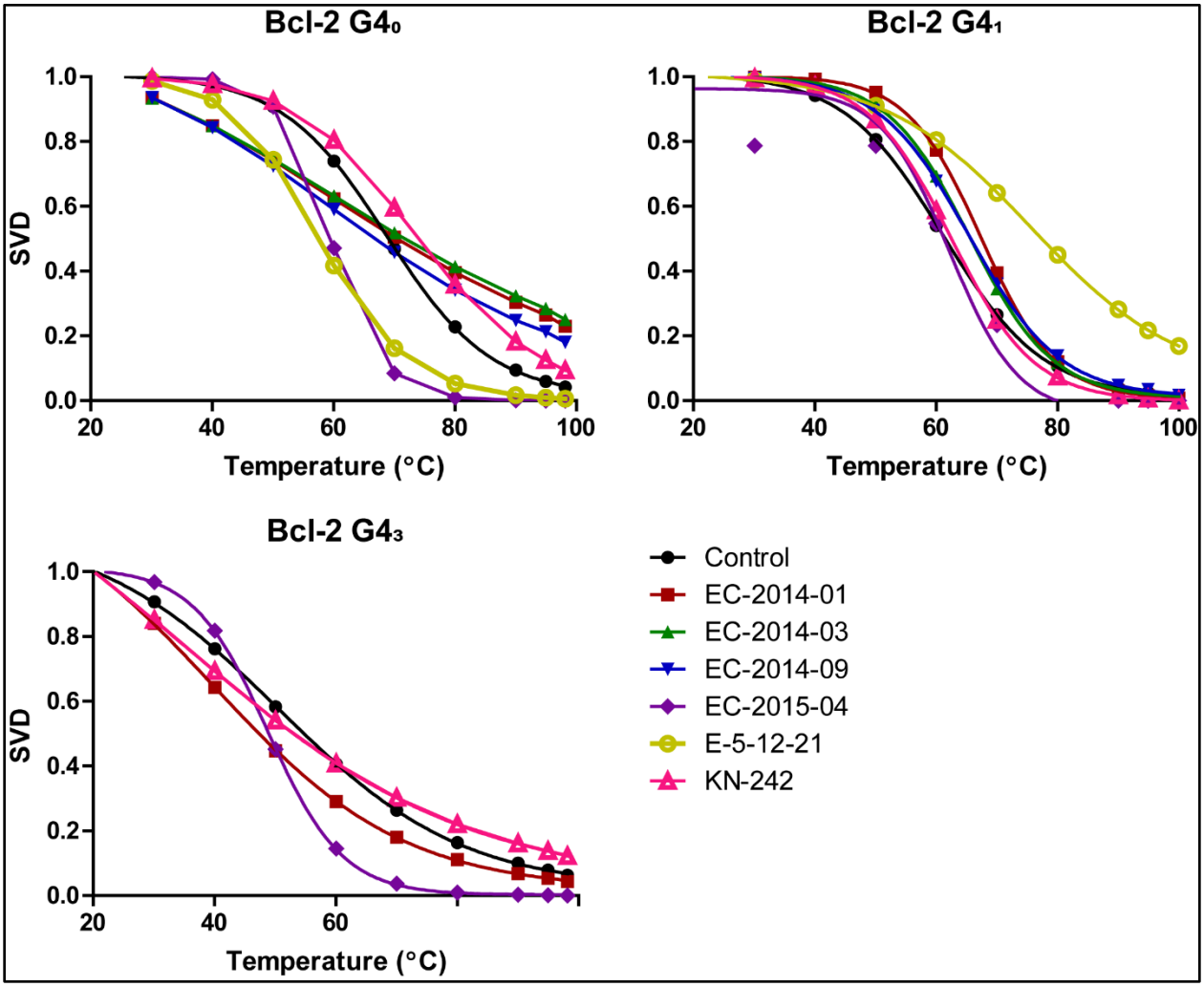


Figure 5-4. FRET melt lead compounds, EC-2014-01 and E-5-12-21, are selective for $G4_1$ stabilization**.** FRET melt hits from the MeCN + KCl plate were confirmed with ECD analysis in multiple Bcl-2 G4s. Transcriptional silencers $G4_0$ and $G4_3$ had decreased thermal stability in the presence of EC-2014-01 and E-5-12-21. $G4_3$ overall formation was inhibited by E-5-12-21, EC-2014-09, and EC-2014-03. KN-242 increased the thermal stability of $G4_0$. $G4_1$ had marked increases in thermal stability upon EC-2014-01 and E-5-12-21 additions. $G4_2$ effects were not examined as the biological role in modulating transcription remains to be elucidated.

5.3. Selective Small Molecule Targeting of $G4_1$ over Other Bcl-2 G4s

Upon ECD validation, compounds EC-2014-01 and E-5-12-21 were tested for any inherent cytotoxicity using the MTS cell viability assay. Neuroblastoma cells, SH-SY5Y, were used as a dopaminergic neuron model like previous reports (Storch et al. 2000; Zhou et al. 2014; Masci et al. 2015). After 72 hr treatment, neither compound showed signs of cytotoxicity up to 100 μM doses (data not shown).

Next, each compounds' impact on promoter activity was evaluated using a Dual-Reporter Luciferase assay where Renilla luciferase expression was driven by pRL (Figure 5-5A). Promoter activity was determined by the firefly to Renilla ratio recorded. The Bcl-2 plasmid used contains the full promoter sequence harboring all four G4-forming regions, and we used promoterless plasmid, empty vector (EV), and promoter-driven plasmid that does not contain a G4-forming region, SV40, as our negative and positive controls, respectively. This assay was performed in both SH-SY5Y and HEK-293 cells. After 24 hr transfection in SH-SY-5Y cells, significant increases in Bcl-2 promoter activity was seen with 1 μM E-5-12-21 (2.04 ± 0.12 fold RLU) compared to both plasmid controls and to its own untreated and DMSO (1.09 ± 0.37 fold RLU) treated cells. Compounds EC-2014-01 (0.48 ± 0.20 fold RLU) and KN-242 (0.55 ± 0.58 fold RLU) showed no marked effect on promoter activity. When performed in HEK-293 cells, marked increases in Bcl-2 promoter activity was seen with KN-242 (2.12 ± 0.07 fold RLU) but not with E-5-12-21 (1.24 ± 0.05 fold RLU) compared to DMSO treated cells (1.05 ± 0.03 fold RLU) (data not shown). This suggests that the transcriptional enhancing effect within the Bcl-2 promoter by E-5-12-21 is specific to neurons.

Compound E-5-12-21 was next tested for having a neuroprotective role against chemical stressor and neurotoxin, 6-hydroxydopamine (6-OHDA) (Figure 5-5C). Using SH-SY5Y cells, neurotoxicity was induced by 24 hr treatment of 100 μM 6-OHDA (Storch et al. 2000). Neuroprotection was tested against this through co-treatment of compound E-5-12-21 (3, 10, and 30 μM doses) for 24 hr with 6-OHDA. Bcl-2 mRNA expression was observed by qRT-PCR. There was no significant change in Bcl-2 expression compared to untreated or DMSO treated cells. This was performed in duplicate, and due to the amount of error, more replicates are ongoing.

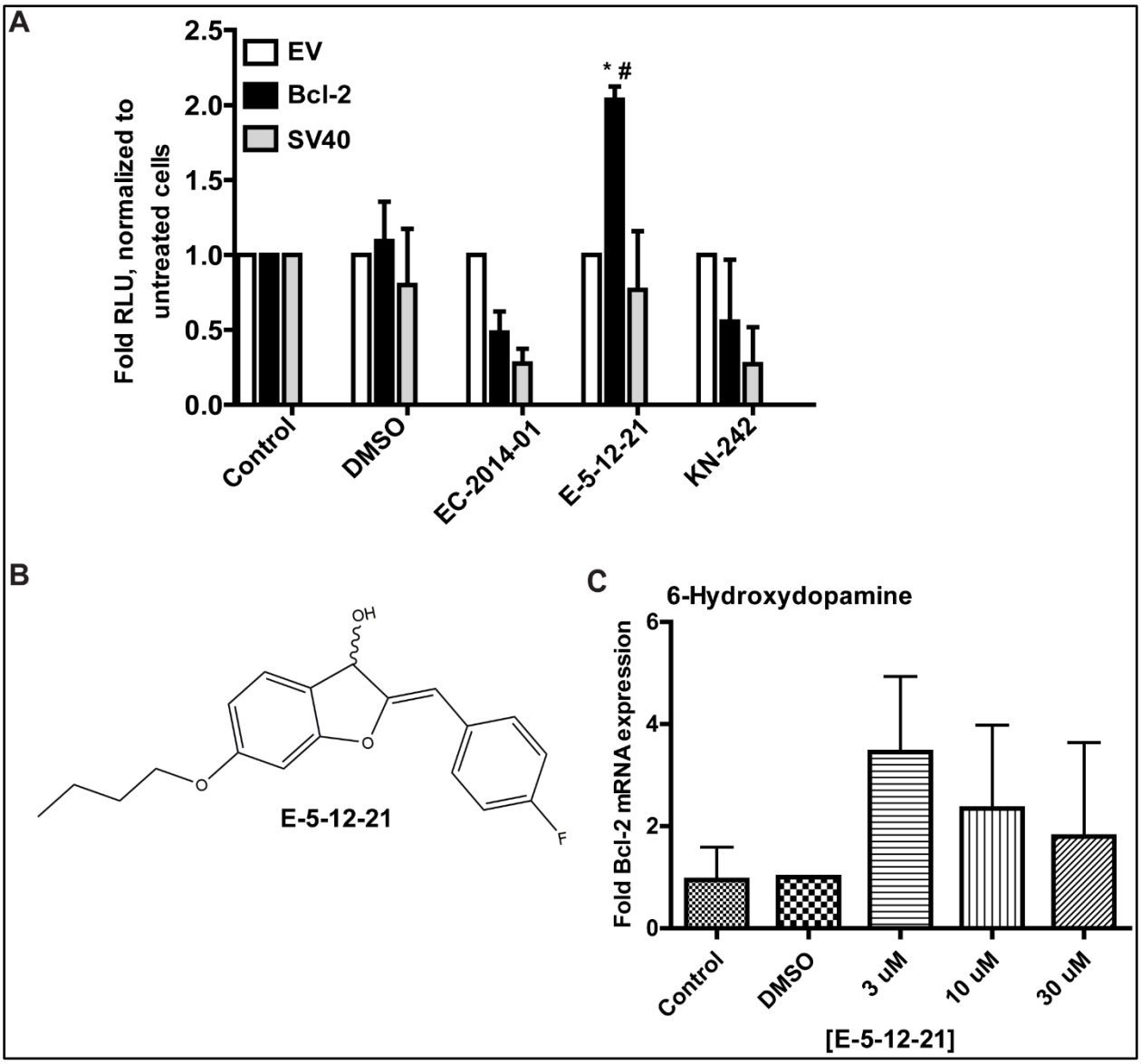


Figure 5-5. Endocannabinoid E-5-12-21 increases Bcl-2 activity. After 24 hr co-transfection and treatment with lead compounds EC-2014-01, E-5-12-21, and KN-242, a significant increase in promoter activity was seen upon E-5-12-21 addition compared to its own plasmids and DMSO treated cells, and no effect was seen with other compounds. Significance was examined by two-way ANOVA with Bonferroni post-test (*$p<0.05$ compared to Bcl-2 control and DMSO, #$p<0.01$ compared to EV and SV40) (A) Endocannabinoid E-5-12-21 structure synthesized by Dr. Eric Bow at the University of Mississippi (B) E-5-12-21 has no significant effect on Bcl-2 mRNA expression under neurotoxic conditions induced by 24 hr treatment of 100 μM 6-hydroxydopamine, performed in biological duplicate (C)

5.4. Discussion

There exist four major G4s that form within the Bcl-2 core promoter encompassed within 500 bp of the TSS. Herein, we focused on targeting $G4_1$ located immediately upstream of the P1 proximal promoter. Though four structures exist, only three have known biological functions on transcription. Silencing occurs most frequently as found throughout G4 research, but $G4_1$ acts as an enhancing element which makes it a novel target for anti-apoptotic-induced neuroprotection.

In this work, we utilized an array of co-solvents to determine optimal concentrations in conjunction with Bcl-2 $G4_1$ oligonucleotides to induce maximal stable structure formation. These experiments allowed us to gain insight into how biomolecules influence G4 dynamics physiologically. From previous work, we observed that co-solute concentrations differ depending on the G4 being examined. For example, in our MYC G4 studies described previously, 20% MeCN was enough to induce G4 formation, but here, for Bcl-2 $G4_1$, we needed 40%. Similar to MYC, this concentration was applied during FRET melt screens for lead compounds stabilizing $G4_1$. We observed an 18% decrease in “hit” compounds with 10 mM KCl + 40% MeCN added compared to the buffer control plate. The same 96-well plate containing compounds was used for each condition, and the leads did not all coincide from plate to plate. ECD later confirmed that two compounds from the MeCN treated FRET melt plate remained to be lead stabilizing agents for Bcl-2 $G4_1$ over the other G4s existing within the same promoter region. Using these compounds to bind and stabilize $G4_1$ is an innovative approach to treating PD.

Current therapeutics focus on attenuating symptoms for patients suffering from PD, as the underlying pathology is unknown. This is particularly disheartening as PD

remains the second most common neurodegenerative disease, Alzheimer's disease being the first, with deaths occurring within ten years following the presence of symptoms. The only pathophysiology known is that dopamine-producing neurons are depleted within the substantia nigra and within the associated nigrostriatal pathway (Katzung et al. 2012). This cell death may be genetically due to mutations in α-synuclein causing the production of Lewy bodies or in parkin resulting in marked protein degradation (Mattson 2000; Angot et al. 2010). On the other hand, dopamine catabolism occurs naturally as adults lose approximately 13% of dopaminergic neurons every ten years. Once 80% of such neurons are depleted, PD symptoms can arise. Lastly, environmental factors and chemicals may induce PD. The best example of a chemical causing PD is MPTP, a precursor to neurotoxin MPP+. This toxin gained much attention in the early 1980s as chronic opioid users looked to abusing MPPP where MPTP was an unfortunate byproduct (Fahn 1996; Youdim & Arraf 2004). MPTP causes permanent PD symptoms and selectively kills dopaminergic neurons within the substantia nigra.

To combat PD, most therapeutics that are prescribed are to increase dopamine (Potenza et al. 2007; Schapira 2007; Allain et al. 2008; Katzung et al. 2012). Sinemet is a common drug given to increase dopamine in the brain and not in the periphery (Hauser et al. 2008). Amantidine allows cells to generate more dopamine by enhancing its release and slowing down reuptake. Another approach is to prevent the destruction of dopamine using monoamine oxidase B (MAO-B) inhibitors like Rasagiline or catechol-oxygen-methyl transferase (COMT) inhibitors like Entacopone (Olanow et al. 2008; Seeberger & Hauser 2009; Youdim et al. 2005). Dopamine receptor agonists are also prescribed to

give something “like dopamine”. Common drugs agonizing the D_2 receptor are Bromocriptine and Ropinirole (Pahwa et al. 2007).

Besides these methods, ulterior ways to manage motor complications through the use of anticholinergics or antihistamines show potential. Lastly, drugs for neuroprotection can be administered that do not target symptomatic relief as these previously described drugs do, instead, neuroprotective agents aim to block the loss of dopaminergic neurons. Ultimately, these agents should slow or even reverse disease progression (Suchowersky et al. 2006). The MAO-B inhibitors mentioned above possess antioxidant, as well as, some anti-apoptotic activity and have shown clinical improvement of the disease over 7 years. Similarly, mitochondrial enhancers, creatine and CoEnzyme Q_{10}, showed improvement if taken early on. However, anti-apoptotic agents should be the sole focus for PD rational drug design as they have a myriad of mechanisms including antioxidant activity and mitochondrial enhancement. Hence, the work presented herein focused on targeting the gene that encodes anti-apoptotic protein Bcl-2.

Our approach to target secondary DNA structures for transcriptional control of the Bcl-2 gene using small molecules is unlike other methods used to treat PD. The lead compound E-5-12-21 identified in a blind small molecule screen is an endocannabinoid synthesized at the University of Mississippi by Dr. Eric Bow. This compound selectively stabilizes $G4_1$ and has no inherent cytotoxicity within dopaminergic neurons. It significantly upregulates Bcl-2 promoter activity in vitro as shown by luciferase. Notably, this increase was only seen within our neuron model and not while using HEK-293 cells. This selectivity of neurons over other cell types shows great promise for this compound

being used as a neuroprotective agent assuming it can easily cross the blood brain barrier (BBB).

Future studies consist of further neuroprotection studies using E-5-12-21 under environmental and inflammatory stressors (serum starvation and H_2O_2, respectively), competition dialysis to determine relative binding affinity for $G4_1$, SPR to reveal dissociation of the compound to our structure of interest, and, lastly, this validated pharmacophore can be administered to PD-induced mice and/or zebrafish for in vivo efficacy and toxicological studies. More FRET melt screens may be performed to yield additional lead compounds. These additional small molecule screens should include the use of nucleoplasm and other key co-solvents like ficoll, dextran sulfate, sucrose, and glucose. Structurally, extensive characterization of the physiologically relevant G4 within the Bcl-2 promoter can be accomplished with plasmid or in cell DMS footprinting. Collectively, these works will positively impact patients suffering from PD and expand current G4 drug discovery efforts from cancer therapeutics to including neuroprotection as well.

CHAPTER 6. CONCLUSION

Non-B-DNA structures, such as G4s, have shown promise in drug discovery. Applying G4 stabilization to transcriptional control of gene regulation in order to suppress oncogene expression is promising for cancer and neuroprotective therapeutic approaches alike. G4-mediated transcriptional silencing or enhancing is generally due to disruption of transcription factor binding due to a change in DNA structure. Stabilization of G4s within promoters act as blockades to RNA polymerase and gene regulatory proteins. Enhancing G4 formation and stability is therefore potentially therapeutic, and as the G4 is more “globular” than traditional dsDNA, small molecules can be more specific for these structures. Extracellularly, stabilization can occur by the addition of alkali cations, modifying water content, or by integrating molecular crowding. Intracellularly, stabilization can be attained pharmacologically by small molecules. Ligand-mediated G4 stabilization has been a growing drug discovery program since the initial G4 characterization in telomeres.

5-hmC epigenetic modifications are largely found to coexist on CpG islands that contain G4-forming regions than its precursor, 5-mC (Bhavsar-Jog, Van Dornshuld, T. a Brooks, et al. 2014). Thus, we looked at the possible role such modifications may play on G4-forming regions by affecting structure formation and stability. Due to its promoter containing CpG islands and the extensive research done on its characterization, the VEGF G4 was used as a model (Sun et al. 2005; Guo et al. 2008; Yoshida et al. 2016).

We found that regardless of the location of the cytosine, or the loop in which it is contained in, there is no significant effect on G4 formation and stability even in the presence of monovalent cations, osmolytes, or molecular crowding agents. Lastly, we observed an increase in G4 stability with epigenetic modifications upon addition of TMPyP4 compared to wild-type.

One question that remains is how these modifications, which reside within G4 loops, influence regulatory protein binding. From there, 5-hmC modifications on the VEGF G4 can be incorporated into a plasmid to monitor the effects on promoter activity compared to wild-type with and without known G4-stabilizing agents. Upon further understanding of epigenetic modifications and G4 formation, stability, and biological function, drug discovery efforts can begin in order to identify compounds that bind and stabilize wild-type and modified VEGF G4s alike resulting in reduced promoter activity and overall inhibition of angiogenesis in tumor cells.

Current drug discovery and development efforts for G4s focus on the use of small molecules. With these programs, several compound classes have been found to stabilize G4s and impact promoter activity in plasmids of the respective gene. However, these lead compounds did not always show activity in cells or, more importantly, in vivo. Only one compound has been introduced into clinical trials. Thus, this field needs better screening methods that coincide with the intracellular conditions of a nucleus to narrow the number of lead compounds that yield higher success rates when transitioning from plasmids into cells, and, ultimately, into animals and humans.

Using the best characterized G4 to date, MYC, we looked at using various co-solvents and extracted nucleoplasm to accomplish mimicking the environment of a

nucleus outside of cells for drug screening. We observed a marked decrease in hit rates (10% reduction) from simple buffer conditions to the addition of KCl and osmolytes MeCN. Some of the differences in hit rates coincide with previous findings on compounds like quindoline i and TMPyP4 (Ou et al. 2007; Boddupally et al. 2012; Izbicka et al. 1999; Siddiqui-Jain et al. 2002). The remaining lead small molecules will be analyzed for anticancer activity in lymphoma cells, and the mechanism of action (G4-dependent or – independent) will be determined by qRT-PCR using the exon test (Boddupally et al. 2012). Lastly, remaining relevant co-solvents will be used in FRET melt screening similar to MeCN shown here. Overall, this will enhance the current methods utilized in drug discovery and development programs targeting G4s, especially those in complex promoters like kRAS and Bcl-2.

kRAS is one of the leading oncogenes in the cancer research field. It is highly deregulated and promotes tumorigenesis in a third of all cancer types, with over 90% frequency in pancreatic ductal adenocarcinomas. Several inhibition approaches have been attempted including disruption of membrane localization, inhibition of downstream effectors, competing with intrinsic GTP, and antisense. Each of these tactics failed, thus, there is a critical need for new targets. We identified a novel molecular target within the guanine dense kRAS core promoter that has the potential to silence transcription resulting in kRAS downregulation and tumor cell death. We conclude that $G4_{mid}$ should be the central focus for rational drug design over its non-biologically active neighbor, $G4_{near}$.

Utilizing NSC 317605, we observed selective stabilization of $G4_{mid}$ over $G4_{near}$ and downregulation of kRAS. Though this compound was not specific for kRAS G4s as it stabilizes MYC as well, synthesizing analogues may yield greater specificity and

selectivity for $G4_{mid}$. Upon identification of $G4_{mid}$-stabilizing agents, anticancer activity and the mechanism of action can be validated in a pancreatic cancer cell line panel, colorectal, and lung cells, as well as, in primary patient biopsies of the same cancer types. These models will represent both mutant forms of kRAS and tumor cells where kRAS is upregulated. From there, murine xenograft models can determine in vivo efficacy of lead pharmacophores and zebrafish for toxicological assessment. Outside of small molecule screening, $G4_{mid}$ remains to be characterized under supercoiled conditions or within cells. Therefore, structurally, $G4_{mid}$ can be elucidated using plasmid and in cell footprinting to determine the biologically relevant G4 isoform (Sun & Hurley 2009; Guo & Bartel 2016). This knowledge will aid in drug discovery efforts using this molecular model and will have a positive impact on cancer (pancreatic, lung, or colorectal) patients whose tumors are influenced by mutant and upregulated kRAS.

Similar to pancreatic cancer, neurodegeneration is a critical issue worldwide with limited therapy options. PD is one of the most common diagnoses each year. Treatments are limited in efficacy due to the unknown etiology of the disease and mechanisms of disease progression are not fully understood. Current therapeutic strategies rely on increasing dopamine levels, activating the motor cortex differently, or neuroprotection. The current work identified a new molecular target for PD resulting in increased promoter activity and anti-apoptosis in neurons. $G4_1$ is one of four G4s within the Bcl-2 core promoter and the only one capable of enhancing transcription making it essential for PD drug design and development.

Using the co-solvent approach derived in our MYC G4-related studies, upon 40% MeCN and 10 mM KCl addition, hit rates dropped 18%. Two compounds were validated

by ECD to selectively stabilize $G4_1$ over the others. With E-5-12-21 being our lead compound, further mechanism of action confirmation is required using a series of neuroprotective studies examining environmental and inflammatory stress with serum starvation and H_2O_2 treatment, respectively. From here, other cannabinoid analogues and natural products can be screened for $G4_1$ stabilization. Lead pharmacophores can be administered to PD murine models for in vivo efficacy assessments. Also, zebrafish models for neurodegeneration can be utilized to validate pharmacophores in vivo. Zebrafish continue to be widely accepted for model organisms in various fields of research as there are over 3,000 human disease genes with zebrafish orthologues (Howe et al. 2013). Also, zebrafish are less expensive compared to rodent models and reproduce rapidly with high numbers of offspring. A PD zebrafish model has already been established using the neurotoxin MPTP (Xi et al. 2011). MPTP-induced PD in zebrafish results in decreased dopamine levels and different behavior patterns (Anichtchik et al. 2004). Zebrafish have an orthologue of the mammalian Bcl-2 protein which can be monitored upon PD-induction (Jette et al. 2008). This molecular model will be one of the first ways to target PD by means of transcriptional control.

Collectively, these works give insight on G4 formation in physiological conditions. Specifically, we examined the role epigenetic modifications play in regulation and how important dehydration and molecular crowding is in promoting structure formation. Poised with this knowledge, G4 drug discovery and development programs can improve the current screening methods used to incorporate physiological conditions. Overall, this will positively impact cancer and neurodegenerative disease patients alike and expand G4

drug discovery efforts to include other critical genes that are significant in biomedical research.

LIST OF REFERENCES

REFERENCES

Adjei, A., 2001. Blocking oncogenic Ras signaling for cancer therapy. *J Natl Cancer Inst*, 93(14), pp.1062–1074.

Agarwala, P. et al., 2013. The G-quadruplex augments translation in the 5' untranslated region of transforming growth factor B2. *Biochemistry*, 52(9), pp.1528–1538.

Agrawal, P. et al., 2013. Solution structure of the major G-quadruplex formed in the human VEGF promoter in K + : insights into loop interactions of the parallel G-quadruplexes. , 41(22), pp.10584–10592.

Ahmed, S., Kintanar, A. & Henderson, E., 1994. Human telomeric C–strand tetraplexes. *Nature Structural & Molecular Biology*, 1(2), pp.83–88.

Akhtar, R.S., Ness, J.M. & Roth, K. a., 2004. Bcl-2 family regulation of neuronal development and neurodegeneration. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research*, 1644(2–3), pp.189–203. Available at: http://linkinghub.elsevier.com/retrieve/pii/S0167488903001848.

Allain, H., Bentué-Ferrer, D. & Akwa, Y., 2008. Disease-modifying drugs and Parkinson's disease. *Progress in Neurobiology*, 84(1), pp.25–39.

Ambrus, A. et al., 2006. Human telomeric sequence forms a hybrid-type intramolecular G-quadruplex structure with mixed parallel/antiparallel strands in potassium solution. *Nucleic Acids Research*, 34(9), pp.2723–2735.

Angot, E. et al., 2010. Are synucleinopathies prion-like disorders? *The Lancet Neurology*, 9(11), pp.1128–1138.

Anichtchik, O. V et al., 2004. Neurochemical and behavioural changes in zebrafish Danio rerio after systemic administration of 6-hydroxydopamine and 1-methyl-4-phenyl-1,2,3,6-tetrahydropyridine. *J. Neurochem*, 88, pp.443–453.

Antequera, F. & Bird, a, 1993. Number of CpG islands and genes in human and mouse. *Proceedings of the National Academy of Sciences of the United States of America*, 90(24), pp.11995–11999.

Aoki, K. et al., 1997. Suppression of Ki-ras p21 levels leading to growth inhibition of pancreatic cancer cell lines with Ki-ras mutation but not those without Ki-ras mutation. *Molecular Carcinogenesis*, 20(2), pp.251–258.

Athuluri-Divakar, S.K. et al., 2016. A Small Molecule RAS-Mimetic Disrupts RAS Association with Effector Proteins to Block Signaling. *Cell*, 165(3), pp.643–655. Available at: http://dx.doi.org/10.1016/j.cell.2016.03.045.

Bailey, P. et al., 2016. Genomic analyses identify molecular subtypes of pancreatic cancer. *Nature*, 531(7592), pp.47–52. Available at: http://www.nature.com/nature/journal/v531/n7592/full/nature16965.html%5Cnhttp://www.nature.com/nature/journal/v531/n7592/pdf/nature16965.pdf.

Balasubramanian, S. & Hurley, L.H., 2011. UKPMC Funders Group Targeting G-quadruplexes in gene promoters : a novel anticancer. , 10(4), pp.261–275.

Bang, I., 1910. Untersuchungen Cuber die GuanylsCaure. *Biochem Z*, 26, pp.293–311.

Bari, L. Di & Pescitelli, G., 2010. Chapter 9: Electronic Circular Dichroism. In *Computational Spectoscopy: Methods, Experiments and Applications*. pp. 241–277.

Bestor, T.H., 2000. The DNA methyltransferases of mammals. *Human molecular genetics*, 9(16), pp.2395–2402.

Bhavsar-Jog, Y.P., Van Dornshuld, E., Brooks, T.A., et al., 2014. Epigenetic modification, dehydration, and molecular crowding effects on the thermodynamics of i-motif structure formation from C-Rich DNA. *Biochemistry*, 53(10), pp.1586–1594.

Bhavsar-Jog, Y.P., Van Dornshuld, E., Brooks, T. a, et al., 2014. Epigenetic modification, dehydration, and molecular crowding effects on the thermodynamics of i-motif structure formation from C-Rich DNA. *Biochemistry*, 53(10), pp.1586–1594.

Bidzinska, J. et al., 2013. G-quadruplex structures in the human genome as novel therapeutic targets. *Molecules*, 18(10), pp.12368–12395.

Biffi, G. et al., 2013. Quantitative Visualization of DNA G-quadruplex Structures in Human Cells. *Nature Chemistry*, 5(3), pp.182–186.

Boddupally, P.V.L. et al., 2012. Anticancer activity and cellular repression of c-MYC by the G-quadruplex-stabilizing 11-piperazinylquindoline is not dependent on direct targeting of the G-quadruplex in the c-MYC promoter. *Journal of Medicinal Chemistry*, 55(13), pp.6076–6086.

Bonnal, S. et al., 2003. A single internal ribosome entry site containing a G quartet RNA structure drives fibroblast growth factor 2 gene expression at four alternative translation initiation codons. *Journal of Biological Chemistry*, 278(41), pp.39330–39336.

Britten, C.D., 2013. PI3K and MEK inhibitor combinations: Examining the evidence in selected tumor types. *Cancer Chemotherapy and Pharmacology*, 71(6), pp.1395–1409.

Brooks, T.A. & Hurley, L.H., 2010. Targeting MYC Expression through G-Quadruplexes. *Genes & cancer*, 1(6), pp.641–649. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2992328&tool=pmcentrez&rendertype=abstract.

Brooks, T.A., Kendrick, S. & Hurley, L., 2010. Making sense of G-quadruplex and i-motif functions in oncogene promoters. *FEBS Journal*, 277(17), pp.3459–3469.

Brooks, T. a & Hurley, L.H., 2009. The role of supercoiling in transcriptional control of MYC and its importance in molecular therapeutics. *Nature reviews. Cancer*, 9(12), pp.849–861.

Brown, R. V. et al., 2011. Demonstration that drug-targeted down-regulation of MYC in non-Hodgkins lymphoma is directly mediated through the promoter G-quadruplex. *Journal of Biological Chemistry*, 286(47), pp.41018–41027.

Bryan, T.M. & Baumann, P., 2011. G-Quadruplexes: From Guanine Gels to Chemotherapeutics. *Molecular Biotechnology*, 49(2), pp.198–208. Available at: http://link.springer.com/10.1007/s12033-011-9395-5.

Burbulla, L.F. et al., 2010. Dissecting the role of the mitochondrial chaperone mortalin in Parkinson's disease: Functional impact of disease-related variants on mitochondrial homeostasis. *Human Molecular Genetics*, 19(22), pp.4437–4452.

Buscaglia, R. et al., 2013. Polyethylene glycol binding alters human telomere G-quadruplex structure by conformational selection. *Nucleic Acids Research*, 41(16), pp.7934–7946. Available at: http://nar.oxfordjournals.org/lookup/doi/10.1093/nar/gkt440.

Carey, N., 2011. *Epigenetics revolution: How modern biology is rewriting our understanding of genetics, disease and inheritance*, London: Icon Books.

Castellano, E. et al., 2013. Requirement for Interaction of PI3-Kinase p110?? with RAS in Lung Tumor Maintenance. *Cancer Cell*, 24(5), pp.617–630.

Chambers, V.S. et al., 2015. High-throughput sequencing of DNA G-quadruplex structures in the human genome. *Nature Biotechnology*, 33(July), pp.1–7. Available at: http://www.nature.com/doifinder/10.1038/nbt.3295 [Accessed August 27, 2015].

Champoux, J.J., 2001. DNA TOPOISOMERASES : Structure, Function, and Mechanism. *Annu. Rev. Biochem.*, 70, pp.369–413.

Chargaff, E., 1950. Chemical specificity of nucleic acids and mechanism of their enzymatic degradation. *Experientia*, 6, pp.201–209.

Chen, Y. & Yang, D., 2013. Sequence, stability, structure of G-quadruplexes and their drug interactions. *Current Protoc Nucleic Acid Chem*, (520), pp.1–26.

Chipuk, J.E. et al., 2010. The BCL-2 Family Reunion. *Molecular Cell*, 37(3), pp.299–310. Available at: http://dx.doi.org/10.1016/j.molcel.2010.01.025.

Cho, K. & Lee, K., 2002. Chemistry and biology of Ras farnesyltransferase. *Arch Pharm Res*, 25(6), pp.759–769.

Chung, H.-J. & Levens, D., 2005. C-Myc Expression: Keep the Noise Down! *Molecules and cells*, 20(2), pp.157–166. Available at: http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=16267388&retmode=ref&cmd=prlinks%5Cnpapers2://publication/uuid/4AEB8F60-7736-488F-B932-0B77769D6604.

Cogoi, S. et al., 2009. Identification of a new G-quadruplex motif in the KRAS promoter and design of pyrene-modified G4-decoys with antiproliferative activity in pancreatic cancer cells. *Journal of medicinal chemistry*, 52(2), pp.564–568.

Cogoi, S. et al., 2013. MAZ-binding G4-decoy with locked nucleic acid and twisted intercalating nucleic acid modifications suppresses KRAS in pancreatic cancer cells and delays tumor growth in mice. *Nucleic Acids Research*, 41(7), pp.4049–4064.

Cogoi, S. et al., 2008. Structural polymorphism within a regulatory element of the human KRAS promoter: Formation of G4-DNA recognized by nuclear proteins. *Nucleic Acids Research*, 36(11), pp.3765–3780.

Cogoi, S. et al., 2010. The KRAS promoter responds to Myc-associated zinc finger and poly(ADP-ribose) polymerase 1 proteins, which recognize a critical quadruplex-forming GA-element. *Journal of Biological Chemistry*, 285(29), pp.22003–22016.

Cogoi, S. & Xodo, L.E., 2006. G-quadruplex formation within the promoter of the KRAS proto-oncogene and its effect on transcription. *Nucleic Acids Research*, 34(9), pp.2536–2549.

Corradini, C.P. et al., 2012. G-Quadruplexes from Human Telomeric DNA: How Many Conformations in PEG Containing Solutions?

Cox, A.D. et al., 2015. HHS Public Access. , 13(11), pp.828–851.

Cui, J. et al., 2013. The effect of molecular crowding on the stability of human c-MYC promoter sequence I-motif at neutral pH. *Molecules*, 18(10), pp.12751–12767.

Dai, J. et al., 2006. NMR solution structure of the major G-quadruplex structure formed in the human BCL2 promoter region. *Nucleic Acids Research*, 34(18), pp.5133–5144. Available at: http://nar.oxfordjournals.org/lookup/doi/10.1093/nar/gkl610.

Dailey, M.M. et al., 2010. Resolution and characterization of the structural polymorphism of a single quadruplex-forming sequence. *Nucleic Acids Research*, 38(14), pp.4877–4888.

Dang, C. V. et al., 2006. The c-Myc target gene network. *Seminars in Cancer Biology*, 16(4), pp.253–264.

Davis-smyth, T., 2016. The Biology of Vascular Endothelial Growth Factor. , 18(1), pp.4–25.

Drygin, D. et al., 2009. Anticancer activity of CX-3543: A direct inhibitor of rRNA biogenesis. *Cancer Research*, 69(19), pp.7653–7661.

Ebi, H. et al., 2011. Receptor tyrosine kinases exert dominant control over PI3K signaling in human KRAS mutant colorectal cancers. *The Journal of Clinical Investigation*, 121(11), pp.4311–4321.

Eddy, J., 2006. Gene function correlates with potential for G4 DNA formation in the human genome. *Nucleic Acids Research*, 34(14), pp.3887–3896. Available at: http://nar.oxfordjournals.org/lookup/doi/10.1093/nar/gkl529.

Eisenman, R.N., 2001. Deconstructing Myc. *Genes and Development*, 15(16), pp.2023–2030.

End, D., 1999. Farnesyl protein transferase inhibitors and other therapies targeting the Ras signal transduction pathway. *Invest New Drugs*, 17(3), pp.241–258.

Engelman, J. a et al., 2008. Effective use of PI3K and MEK inhibitors to treat mutant Kras G12D and PIK3CA H1047R murine lung cancers. *Nature medicine*, 14(12), pp.1351–6. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2683415&tool=pmcentrez&rendertype=abstract.

Fahn, S., 1996. Book Review-The Case of the Frozen Addicts: How the Solution of an Extraordinary Medical Mystery Spawned a Revolutino in the Understanding and Treatment of Parkinson's Disease. *The New England Journal of Medicine*, 335(26), pp.2002–2003.

Farlie, P.G. et al., 1995. Bcl-2 Transgene Expression Can Protect Neurons Against Developmental and Induced Cell Death. *Proceedings of the National Academy of Sciences of the United States of America*, 92(10), pp.4397–4401. Available at: http://www.pnas.org/content/92/10/4397.long%5Cnpapers2://publication/uuid/1ADB250F-1FF7-4DC7-9616-B04FE6EC6690.

Fossum, B. et al., 1995. CD8+ T cells from a patient with colon carcinoma, specific for a mutant p21-Ras-derived peptide (Gly13-->Asp), are cytotoxic towards a carcinoma cell line harbouring the same mutation. *Cancer Immunol Immunother*, 40(3), pp.165–172.

Gellert, I., Lipsett, M.N. & Davies, D.R., 1962. Helix formation by guanylic acid. *Proceedings of the National Academy of Sciences of the United States of America*, 48, pp.2013–2018.

Ghobrial, I. & Adjei, A., 2002. Inhibitors of the ras oncogene as therapeutic targets. *Hematol Oncol Clin North Am*, 16(5), pp.1065–1088.

Gjertsen, M.K. et al., 2001. Intradermal ras peptide vaccination with granulocyte-macrophage colony- stimulating factor as adjuvant: Clinical and immunological responses in patients with pancreatic adenocarcinoma. *International journal of cancer. Journal international du cancer*, 92(3), p.441–50. Available at: http://www.ncbi.nlm.nih.gov/htbin-post/Entrez/query?db=m&form=6&dopt=r&uid=11291084.

Gjertsen, M.K. & Gaudernack, G., 1998. Mutated Ras peptides as vaccines in immunotherapy of cancer. *Vox sanguinis*, 74 Suppl 2, pp.489–95. Available at: http://www.ncbi.nlm.nih.gov/pubmed/9704487.

Goobes, R. et al., 2003. Metabolic buffering exerted by macromolecular crowding on DNA-DNA interactions: Origin and physiological significance. *Biochemistry*, 42(8), pp.2431–2440.

Gray, G.D. et al., 1993. Antisense DNA Inhibition of Tumor Growth Induced by c-Ha-ras

Oncogene in Nude Mice. *Cancer Research*, 53(3), pp.577–580.

Guo, J.U. & Bartel, D.P., 2016. RNA G-quadruplexes are globally unfolded in eukaryotic cells and depleted in bacteria. *Science (New York, N.Y.)*, 353(6306), p.aaf5371-aaf5371. Available at: http://www.sciencemag.org/cgi/doi/10.1126/science.aaf5371%5Cnhttp://www.ncbi.nlm.nih.gov/pubmed/27708011.

Guo, K. et al., 2008. Intramolecularly folded G-quadruplex and i-motif structures in the proximal promoter of the vascular endothelial growth factor gene. , 36(14), pp.4598–4608.

Gupta, S. et al., 2007. Binding of Ras to Phosphoinositide 3-Kinase p110?? Is Required for Ras- Driven Tumorigenesis in Mice. *Cell*, 129(5), pp.957–968.

Guschlbauer, W., Chantot, J. & Thiele, D., 1990. Four-stranded nucleic acid structures 25 years later: from guanosine gels to telomer DNA. *Journal of Biomolecular Structure and Dynamics*, 8(3), pp.491–511.

Halder, R. et al., 2010. Guanine quadruplex DNA structure restricts methylation of CpG dinucleotides genome-wide w. , pp.2439–2447.

Hanahan, D. & Weinberg, R.A., 2011. Hallmarks of cancer: The next generation. *Cell*, 144(5), pp.646–674. Available at: http://dx.doi.org/10.1016/j.cell.2011.02.013.

Hanahan, D. & Weinberg, R.A., 2000. The hallmarks of cancer. *Cell*, 100(1), pp.57–70. Available at: http://www.ncbi.nlm.nih.gov/pubmed/10647931.

Hauser, R. et al., 2008. Double-blind trial of levodopa/carbidopa/entacapone versus levodopa/carbidopa in early Parkinson's disease. *Movement Disorders*, 24(4), pp.541–550.

Heckman, C. et al., 1997. The WT1 protein is a negative regulator of the normal bcl-2 allele in t(14;18) lymphomas. *Journal of Biological Chemistry*, 272(31), pp.19609–19614.

van der Heide, L.P. & Smidt, M.P., 2013. The BCL2 code to dopaminergic development and Parkinson's disease. *Trends in Molecular Medicine*, 19(4), pp.211–216. Available at: http://linkinghub.elsevier.com/retrieve/pii/S1471491413000324.

Henderson, E.R., Moore, M. & Malcolm, B. a, 1990. Telomere G-strand structure and function analyzed by chemical protection, base analogue substitution, and utilization by telomerase in vitro. *Biochemistry*, 29(3), pp.732–737. Available at: http://www.ncbi.nlm.nih.gov/pubmed/2337592.

Hetts, S.W., 1998. To die or not to die: an overview of apoptosis and its role in disease. *JAMA : the journal of the American Medical Association*, 279(4), pp.300–7. Available at: http://www.ncbi.nlm.nih.gov/pubmed/9450715.

Hoffman, A.R. & Hu, J.F., 2006. Directing DNA methylation to inhibit gene expression. *Cell Mol Neurobiol*, 26(4–6), pp.425–438. Available at: http://www.ncbi.nlm.nih.gov/pubmed/16710755.

Howe, K. et al., 2013. The zebrafish reference genome sequence and its relationship to the human genome. *Nature*, 496(7446), pp.498–503. Available at: http://www.nature.com/nature/journal/vaop/ncurrent/full/nature12111.html?utm_source=feedly.

Huang, W. et al., 2012. Yin Yang 1 contains G-quadruplex structures in its promoter and 5′-UTR and its expression is modulated by G4 resolvase 1. *Nucleic Acids Research*, 40(3), pp.1033–1049.

Huppert, J.L., 2010. Structure, location and interactions of G-quadruplexes. *FEBS Journal*, 277(17), pp.3452–3458.

Huppert, J.L. & Balasubramanian, S., 2006. G-quadruplexes in promoters throughout the human genome. *Nucleic Acids Research*, 35(2), pp.406–413. Available at: http://nar.oxfordjournals.org/lookup/doi/10.1093/nar/gkl1057.

Izbicka, E. et al., 1999. Effects of Cationic Porphyrins as G-Quadruplex Interactive Agents in Human Tumor Cells 1. *CANCER RESEARCH*, 59, pp.639–644.

Jabbari, K. & Bernardi, G., 2004. Cytosine methylation and CpG, TpG (CpA) and TpA frequencies. *Gene*, 333, pp.143–149.

Jean-Claude Martinou, Michel Dubois-Dauphin, Julie K. Staple, Ivan Rodriguez, Harald Frankowski, Marc Missotten, Paola Albertini, Dominique Talabot, Stefan Catsicas, Claudio Pietra, J.H., 1994. Overexpression of BCL-2 in transgenic mice protects neurons from naturally occurring cell death and experimental ischemia. *Neuron*, 13(4), pp.1017–1030.

Jette, C. a et al., 2008. BIM and other BCL-2 family proteins exhibit cross-species conservation of function between zebrafish and mammals. *Cell death and differentiation*, 15(6), pp.1063–1072. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3212414&tool=pmcentrez&rendertype=abstract%5Cnhttp://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3212414%7B&%7Dtool=pmcentrez%7B&%7Drendertype=abstract.

Jordano, J. & Perucho, M., 1988. Initial characterization of a potential transcriptional enhancer for the human c-K-ras gene. *Oncogene*, 2(4), pp.359–366. Available at: http://www.ncbi.nlm.nih.gov/htbin-post/Entrez/query?db=m&form=6&dopt=r&uid=3283654.

Josse, J., Kaiser, A.D. & Kornberg, A., 961. Enzymatic synthesis of deoxyribonucleic acid. *Journal of Biological Chemistry1*, 236.3, pp.864–875.

Katzung, B.G., Masters, S.B. & Trevor, A.J., 2012. *Basic & Clinical Pharmacology*,

Kim, B.G. et al., 2016. Ionic Effects on VEGF G-quadruplex Stability. *The journal of physical chemistry. B*. Available at: http://www.ncbi.nlm.nih.gov/pubmed/27196695.

Kim, J.Y. et al., 2009. The expression of VEGF receptor genes is concurrently influenced by epigenetic gene silencing of the genes and VEGF activation. *Epigenetics : official journal of the DNA Methylation Society*, 4(5), pp.313–321.

Kim, N. et al., 1994. Specific association of human telomerase activity with immortal cells and cancer. *Science*, 266(5193), pp.2011–2015.

Kramer, P.R. & Sinden, R.R., 1997. Measurement of unrestrained negative supercoiling and topological domain size in living human cells. *Biochemistry*, 36(11), pp.3151–3158.

Lagnado, J., 2013. The story of quadruplex DNA – it started with a Bang ! *Biochemical Society*, 35(2), pp.44–46.

Lavrado, J. et al., 2015. KRAS oncogene repression in colon cancer cell lines by G-quadruplex binding indolo[3,2-c]quinolines. *Scientific Reports*, 5, p.9696. Available at: http://www.nature.com/doifinder/10.1038/srep09696.

Lieber, M. et al., 1975. Establishment of a continuous tumor-cell line (panc-1) from a human carcinoma of the exocrine pancreas. *International Journal of Cancer*, 15(5), pp.741–747.

Lin, J. et al., 2013. Stabilization of G-quadruplex DNA by C-5-methyl-cytosine in bcl-2 promoter: implications for epigenetic regulation. *Biochemical and biophysical research communications*, 433(4), pp.368–73. Available at: http://www.ncbi.nlm.nih.gov/pubmed/23261425.

Lipps, H.J. & Rhodes, D., 2009. G-quadruplex structures: in vivo evidence and function. *Trends in Cell Biology*, 19(8), pp.414–422. Available at: http://linkinghub.elsevier.com/retrieve/pii/S0962892409001238.

Lowy, D.R. & Willumsen, B.M., 1993. Function and Regulation of RAS. *Annual Review of Biochemistry*, 62, pp.851–891.

Marušič, M. et al., 2013. G-rich VEGF aptamer with locked and unlocked nucleic acid modifications exhibits a unique G-quadruplex fold. *Nucleic Acids Research*, 41(20), pp.9524–9536.

Masci, A. et al., 2015. Neuroprotective Effect of *Brassica oleracea* Sprouts Crude Juice in a Cellular Model of Alzheimer's Disease. *Oxidative Medicine and Cellular Longevity*, 2015, pp.1–17. Available at: http://www.hindawi.com/journals/omcl/2015/781938/.

Matsuo, Y. & Yamazaki, T., 1989. Nucleic Acids Research. *Nucleic Acids Research*, 17(1), pp.1–2.

Mattson, M.P., 2000. Apoptosis in neurodegenerative disorders. *Nature reviews. Molecular cell biology*, 1(2), pp.120–9. Available at: http://www.ncbi.nlm.nih.gov/pubmed/17034345.

Michaelidis, T.M. et al., 1996. Inactivation of bcl-2 results in progressive degeneration of motoneurons, sympathetic and sensory neurons during early postnatal development. *Neuron*, 17(1), pp.75–89.

Michelotti, Gregory A., et. al., 1996. Multiple single-stranded cis elements are associated with activated chromatin of the human c-myc gene in vivo. *Molecular and cellular*

biology, 16(6), p.2656.

Miescher, F., 1871. Ueber die chemische zusammensetzung der eiterzellen. *Medicinisch-Chemische Untersuchungen*, 4, pp.441–460.

Miller, M.C. et al., 2010. Hydration is a major determinant of the G-quadruplex stability and conformation of the human telomere 3' sequence of d(AG 3(TTAG3)3). *Journal of the American Chemical Society*, 132(48), pp.17105–17107.

Misra, V.K. & Draper, D.E., 1998. On the Role of Magnesium Ions in RNA Stability. *Biopolymers*, 48, pp.113–135.

Miyoshi, D., Karimata, H. & Sugimoto, N., 2006. Hydration regulates thermodynamics of G-quadruplex formation under molecular crowding conditions. *Journal of the American Chemical Society*, 128(24), pp.7957–7963.

Morgan, R.K. et al., 2016. Identification and characterization of a new G-quadruplex forming region within the kRAS promoter as a transcriptional regulator. *Biochimica et Biophysica Acta - Gene Regulatory Mechanisms*, 1859(2), pp.235–245. Available at: http://dx.doi.org/10.1016/j.bbagrm.2015.11.004.

Morgan, R.K. & Brooks, T.A., 2016. Chapter 3: Targeting promoter G-quadruplexes for transcriptional control. In *Small-Molecule Transcription Factor Inhibitors in Oncology*.

Morrison, J. M., et. al., 1967. Nearest neighbour base sequence analysis of the deoxyribonucleic acids of a further three mammalian viruses: simian virus 40, human papilloma virus and adenovirus type 2. *Journal of General Virology*, 1.1, pp.101–108.

Murat, P. & Balasubramanian, S., 2014. Existence and consequences of G-quadruplex structures in DNA. *Current Opinion in Genetics & Development*, 25, pp.22–29. Available at: http://linkinghub.elsevier.com/retrieve/pii/S0959437X13001494.

Neidle, S., 2010. Human telomeric G-quadruplex: The current status of telomeric G-quadruplexes as therapeutic targets in human cancer. *FEBS Journal*, 277(5), pp.1118–1125.

Neidle, S., 2009. The structures of quadruplex nucleic acids and their drug complexes. *Current Opinion in Structural Biology*, 19(3), pp.239–250. Available at: http://linkinghub.elsevier.com/retrieve/pii/S0959440X09000463.

Neidle, S. & Parkinson, G.N., 2003. The structure of telomeric DNA. *Current Opinion in Structural Biology*, 13(3), pp.275–283.

Olanow, C. et al., 2008. A randomized, double-blind, placebo-controlled, delayed start study to assess rasagiline as a disease modifying therapy in Parkinson's disease (the ADAGIO study): rationale, design, and baseline characteristics. *Movement Disorders*, 23(15), pp.2194–2201.

Onel, B. et al., 2016. A New G-Quadruplex with Hairpin Loop Immediately Upstream of the Human BCL2 P1 Promoter Modulates Transcription. *Journal of the American Chemical Society*, 138(8), pp.2563–2570.

Onyshchenko, M.I. et al., 2009. Stabilization of G-quadruplex in the BCL2 promoter region in double-stranded DNA by invading short PNAs. *Nucleic Acids Research*, 37(22), pp.7570–7580. Available at: http://nar.oxfordjournals.org/lookup/doi/10.1093/nar/gkp840.

Ostrem, J. & Shokat, K., 2016. Direct small-molecule inhibitors of K-Ras: from structural insights to mechanism-based design. *Nature Publishing Group*, 15(11), pp.771–785. Available at: http://dx.doi.org/10.1038/nrd.2016.139.

Ostrem, J.M. et al., 2013. K-Ras(G12C) inhibitors allosterically control GTP affinity and effector interactions. *Nature*, 503(7477), pp.548–51. Available at: http://dx.doi.org/10.1038/nature12796.

Ou, T.M. et al., 2007. Stabilization of G-quadruplex DNA and down-regulation of oncogene c-myc by quindoline derivatives. *Journal of Medicinal Chemistry*, 50(7), pp.1465–1474.

Paeschke, K. et al., 2008. Telomerase recruitment by the telomere end binding protein-b facilitates G-quadruplex DNA unfolding in ciliates. *Nature Structural & Molecular Biology*, 15, pp.598–604.

Pahwa, R. et al., 2007. Ropinirole 24-hour prolonged release: randomized, controlled study in advanced Parkinson disease. *Neurology*, 68(14), pp.1008–1115.

Paramasivam, M. et al., 2009. Protein hnRNP A1 and its derivative Up1 unfold quadruplex DNA in the human KRAS promoter: Implications for transcription. *Nucleic Acids Research*, 37(9), pp.2841–2853.

Paramasivam, M., Cogoi, S. & Xodo, L.E., 2011. Primer extension reactions as a tool to uncover folding motifs within complex G-rich sequences: analysis of the human KRAS NHE. *Chemical communications (Cambridge, England)*, 47(17), pp.4965–7. Available at: http://www.ncbi.nlm.nih.gov/pubmed/21424008.

Paramasivan, S., Rujan, I. & Bolton, P.H., 2007. Circular dichroism of quadruplex DNAs: Applications to structure, cation effects and ligand binding. *Methods*, 43(4), pp.324–331. Available at: http://linkinghub.elsevier.com/retrieve/pii/S1046202307000321.

Phan, A.T., Modi, Y.S. & Patel, D.J., 2004. Propeller-type parallel-stranded G-quadruplexes in the human c-myc promoter. *Journal of the American Chemical Society*, 126(28), pp.8710–8716.

Potenza, M.N., Voon, V. & Weintraub, D., 2007. Drug Insight: impulse control disorders and dopamine therapies in Parkinson's disease. *Nature Reviews Neurology*, 3, pp.664–672.

Pray, L.A., 2008. Discovery of DNA Structure and Function: Watson and Crick. *Nature Education*, 1(1), p.100.

Queneau, P. et al., 2001. Early detection of pancreatic cancer in patients with chronic pancreatitis: diagnostic utility of a K-ras point mutation in the pancreatic juice. *Am J Gastroenterol*, 96(3), pp.700–704.

Rangan, A., Fedoroff, O.Y. & Hurley, L.H., 2001. Induction of duplex to G-quadruplex transition in the c-myc promoter region by a small molecule. *Journal of Biological Chemistry*, 276(7), pp.4640–4646.

Rangasamy, S.B. et al., 2010. Neurotrophic factor therapy for Parkinson's disease. *Progress in Brain Research*, 184, pp.237–264.

Rowinsky, E.K., Windle, J.J. & Von Hoff, D.D., 1999. Ras protein farnesyltransferase: A strategic target for anticancer therapeutic development. *Journal of Clinical Oncology*, 17(11), pp.3631–3652.

Saad, I.I., Saha, S.B. & Thomas, G., 2014. The RAS subfamily Evolution - tracing evolution for its utmost exploitation. *Bioinformation*, 10(5), pp.293–8. Available at: http://www.ncbi.nlm.nih.gov/pubmed/24966537%5Cnhttp://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC4070039.

Scarano, E., et. al., 1967. The heterogeneity of thymine methyl group origin in DNA pyrimidine isostichs of developing sea urchin embryos. *Proceedings of the National Academy of Sciences*, 57.5, pp.1394–1400.

Schapira, A.H., 2007. Treatment options in the modern management of Parkinson disease. *Arch Neurol*, 64(8), pp.1083–1088. Available at: http://www.ncbi.nlm.nih.gov/pubmed/17698697.

Seeberger, L.C. & Hauser, R.A., 2009. Levodopa/carbidopa/entacapone in Parkinson's disease. *Expert Review of Neurotherapeutics*, 9(7), pp.929–940.

Seenisamy, J. et al., 2004. The dynamic character of the G-quadruplex element in the c-MYC promoter and modification by TMPyP4. *Journal of the American Chemical Society*, 126(28), pp.8702–8709.

Senger, Donald R., et. al., 1983. Tumor cells secrete a vascular permeability factor that promotes accumulation of ascites fluid. *Science*, 219(4587), pp.983–985.

Shacka, J. & Roth, K., 2005. Regulation of neuronal cell death and neurodegeneration by members of the Bcl-2 family: therapeutic implications. *Current Drug Targets CNS Neurological Disorders*, 4(1), pp.25–39.

Sharma, S., Kelly, T.K. & Jones, P.A., 2009. Epigenetics in cancer. *Carcinogenesis*, 31(1), pp.27–36.

Sharp, K. a., 2016. Unpacking the origins of in-cell crowding. *Proceedings of the National Academy of Sciences*, 113(7), p.201600098. Available at: http://www.pnas.org/lookup/doi/10.1073/pnas.1600098113.

Siddiqui-Jain, A. et al., 2002. Direct evidence for a G-quadruplex in a promoter region and its targeting with a small molecule to repress c-MYC transcription. *Proc. Natl. Acad. Sci. U. S. A.*, 99(18), pp.11593–8. Available at: http://www.pnas.org/content/99/18/11593.full.

Simmons, D., 2008. Epigenetic influence and disease. *Nature Education*, 1(1), p.6.

Simonsson, T., Pribylova, M. & Vorlickova, M., 2000. A nuclease hypersensitive element in the human c-myc promoter adopts several distinct i-tetraplex structures. *Biochemical and biophysical research communications*, 278(1), pp.158–166.

Sissi, C., Gatto, B. & Palumbo, M., 2011. The evolving world of protein-G-quadruplex recognition: A medicinal chemist's perspective. *Biochimie*, 93(8), pp.1219–1230. Available at: http://linkinghub.elsevier.com/retrieve/pii/S0300908411001374.

Song, I.S. et al., 2009. Human ZNF312b promotes the progression of gastric cancer by transcriptional activation of the K-ras gene. *Cancer Research*, 69(7), pp.3131–3139.

Spink, C.H. & Chaires, J.B., 1999. Effects of hydration, ion release, and excluded volume on the melting of triplex and duplex DNA. *Biochemistry*, 38(1), pp.496–508.

Storch, a et al., 2000. 6-Hydroxydopamine toxicity towards human SH-SY5Y dopaminergic neuroblastoma cells: independent of mitochondrial energy metabolism. *Journal of neural transmission (Vienna, Austria : 1996)*, 107(3), pp.281–293.

Stroud, H. et al., 2011. 5-Hydroxymethylcytosine is associated with enhancers and gene bodies in human embryonic stem cells. *Genome biology*, 12(6), p.R54. Available at: http://genomebiology.biomedcentral.com/articles/10.1186/gb-2011-12-6-r54.

Suchowersky, O. et al., 2006. Practice Parameter : Neuroprotective strategies and alternative therapies for Parkinson disease (an evidence-based review) Report of the Quality Standards Subcommittee of the.

Sun, D. et al., 2005. Facilitation of a structural transition in the polypurine / polypyrimidine tract within the proximal promoter region of the human VEGF gene by the presence of potassium and G-quadruplex-interactive agents. , 33(18), pp.6070–6080.

Sun, D. et al., 2008. Proximal promoter region of the human vascular endothelial growth factor gene has a G-quadruplex structure which can be targeted by G-quadruplex-interactive agents. *Molecular Cancer Therapeutics*, 7(4), pp.880–889.

Sun, D., Guo, K. & Shin, Y.-J., 2011. Evidence of the formation of G-quadruplex structures in the promoter region of the human vascular endothelial growth factor gene. *Nucleic acids research*, 39(4), pp.1256–1265.

Sun, D. & Hurley, L.H., 2009. The importance of negative superhelicity in inducing the formation of G-quadruplex and i-motif structures in the c-Myc promoter: implications for drug targeting and control of gene expression. *Journal of medicinal chemistry*, 52(9), pp.2863–74. Available at: http://dx.doi.org/10.1021/jm900055s.

Sun, H. et al., 2014. A newly identified G-quadruplex as a potential target regulating Bcl-2 expression. *Biochimica et Biophysica Acta (BBA) - General Subjects*, 1840(10), pp.3052–3057. Available at: http://linkinghub.elsevier.com/retrieve/pii/S0304416514002591.

Sun, Q. et al., 2012. Discovery of small molecules that bind to K-Ras and inhibit Sos-mediated activation. *Angewandte Chemie - International Edition*, 51(25), pp.6140–6143.

Taberlay, P.C. & Jones, P.A., 2011. DNA methylation and cancer. *Progress in Drug Research*, 67, pp.1–23.

Tsujimoto, Y. et al., 1985. Involvement of the bcl-2 gene in human follicular lymphoma. *Science*, 228(4706), pp.1440–1443.

Turunen, M.P. & Ylä-Herttuala, S., 2011. Epigenetic regulation of key vascular genes and growth factors. *Cardiovascular Research*, 90(3), pp.441–446.

V??h??tupa, M. et al., 2016. Lack of R-Ras leads to increased vascular permeability in ischemic retinopathy. *Investigative Ophthalmology and Visual Science*, 57(11), pp.4898–4909.

Verma, A. et al., 2008. Genome-wide computational and expression analyses reveal G-quadruplex DNA motifs as conserved cis-regulatory elements in human and related species. *Journal of Medicinal Chemistry*, 51(18), pp.5641–5649.

Wang, X.-D. et al., 2010. Turning off Transcription of the *bcl-2* Gene by Stabilizing the *bcl-2* Promoter Quadruplex with Quindoline Derivatives. *Journal of Medicinal Chemistry*, 53(11), pp.4390–4398. Available at: http://pubs.acs.org/doi/abs/10.1021/jm100445e.

Wang, Z. & Dröge, P., 1996. Differential control of transcription-induced and overall DNA supercoiling by eukaryotic topoisomerases in vitro. *The EMBO journal*, 15(3), pp.581–589.

Ward, K.R. et al., 2004. Expression of activated M-Ras in a murine mammary epithelial cell line induces epithelial-mesenchymal transition and tumorigenesis. *Oncogene*, 23(6), pp.1187–1196.

Watson, J.D. & Crick, F.H.C., 1953. A structure for deoxyribose nucleic acid. *Nature*, 171, pp.737–738.

Williamson, J., 1994. G-quartet structures in telomeric DNA. *Annual review of biophysics and biomolecular structure*, 23, pp.703–730.

Wilson, C.Y. & Tolias, P., 2016. Recent advances in cancer drug discovery targeting RAS. *Drug Discovery Today*, 0(0), pp.10–14. Available at: http://linkinghub.elsevier.com/retrieve/pii/S1359644616302896.

Winn, M.D. et al., 2009. NIH Public Access. *Journal of Biomolecular NMR*, 51(4), pp.4108–4139. Available at: http://www.scopus.com/inward/record.url?eid=2-s2.0-0013498057&partnerID=40&md5=7fa5981e2dbf6fccbf58a943e31838a7%5Cnhttp://www.ncbi.nlm.nih.gov/pubmed/24356774%5Cnhttp://www.ncbi.nlm.nih.gov/pubmed/17406304%5Cnhttp://www.nature.com/doifinder/10.1038/nprot.2006.1.

Wolf, G., 2003. Friedrich Miescher: The man who discovered DNA. *Chemical Heritage*, 21, pp.10–11, 37–41.

Xi, Y., Noble, S. & Ekker, M., 2011. Modeling Neurodegeneration in Zebrafish. *Current Neurology and Neuroscience Reports*, 11(3), pp.274–282. Available at: http://link.springer.com/10.1007/s11910-011-0182-2.

Xu, H. et al., 2017. CX-5461 is a DNA G-quadruplex stabilizer with selective lethality in BRCA1/2 deficient tumours. *Nature Communications*, 8(205), p.14432. Available at: http://www.nature.com/doifinder/10.1038/ncomms14432.

Xu, Y. et al., 2010. Telomeric repeat-containing RNA structure in living cells. *Proceedings of the National Academy of Sciences*, 107(33), pp.14579–14584. Available at: http://www.pnas.org/cgi/doi/10.1073/pnas.1001177107.

Yamamoto, F. & Perucho, M., 1988. Characterization of the Human c-K-ras gene Promoter. *Oncogene Research*, 3(2), pp.125–130.

Yoshida, W. et al., 2016. Detection of DNA Methylation of G-Quadruplex and i-Motif-Forming Sequences by Measuring the Initial Elongation Efficiency of Polymerase Chain Reaction. *Analytical Chemistry*, 88(14), pp.7101–7107. Available at: http://pubs.acs.org/doi/abs/10.1021/acs.analchem.6b00982.

Youdim, M.B.H. et al., 2005. Rasagiline: Neurodegeneration, neuroprotection, and mitochondrial permeability transition. *Journal of Neuroscience Research*, 79(1–2), pp.172–179. Available at: http://doi.wiley.com/10.1002/jnr.20350.

Youdim, M.B.H. & Arraf, Z., 2004. Prevention of MPTP (N-methyl-4-phenyl-1,2,3,6-tetrahydropyridine) dopaminergic neurotoxicity in mice by chronic lithium: involvements of Bcl-2 and Bax. *Neuropharmacology*, 46(8), pp.1130–1140.

Zhang, D. et al., 2010. Monomorphic RNA G-Quadruplex and Polymorphic DNA G-Quadruplex Structures Responding to Cellular Environmental Factors †. , pp.4554–4563.

Zhang, Y. et al., 1993. Retroviral Vector-Mediated Transduction of K-ras Antisense RNA into Human Lung Cancer Cells Inhibits Expression of the Malignant Phenotype. *Human Gene Therapy*, 4(4), pp.451–460.

Zhou, J. et al., 2016. Reevaluation of the stability of G-quadruplex structures under crowding conditions. *Biochimie*, 121, pp.204–208. Available at: http://dx.doi.org/10.1016/j.biochi.2015.12.012.

Zhou, M. et al., 2014. Cold-inducible RNA-binding protein mediates neuroinflammation in cerebral ischemia. *Biochimica et Biophysica Acta - General Subjects*, 1840(7), pp.2253–2261. Available at: http://dx.doi.org/10.1016/j.bbagen.2014.02.027.

VITA

Rhianna Kyelle Morgan was born in Hamilton, New York on May 25, 1992. She graduated from Sherburne-Earlville High School (Sherburne, NY) in June 2010. The following fall she entered Hartwick College (Oneonta, NY) where she majored in Mathematics and minored in Biology. After completion of Hartwick's Three Year Degree program in May 2013, she graduated and entered The University of Mississippi to study Pharmacology. Rhianna earned doctoral candidacy in September 2015, and in March 2017, she received her Ph.D. in Pharmaceutical Sciences (pharmacology emphasis).